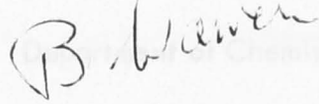
I CERTIFY THAT THIS THESIS IS MY OWN ORIGINAL WORK, EXCEPT WHERE DUE REFERENCE IS MADE TO THE WORK OF OTHERS



THE SEQUENCE DETERMINATION OF PEPTIDES BY NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

A thesis submitted for the degree of

Doctor of Philosophy

in the

Australian National University

by

Brian Warren

Department of Chemistry

Australian National University

November 1976

ACKNOWLEDGEMENTS

I would like to thank the Australian National University for the opportunity to undertake this project by providing a scholarship and travelling grant.

Thanks must also go to my Supervisor, Dr. J.H. Bradbury for his continued encouragement, guidance and endless enthusiasm about this project. It has been a most rewarding experience working with him during the last three years.

I would like to acknowledge the help given by other staff of the Chemistry Department; in particular Mr. R.J. Nicholls for maintaining the NMR spectrometer, Mr. R. Briggs for running mass spectra, and the National NMR Centre for providing 270 MHz spectra.

I am grateful to the staff of H.B. Selby and Co. Ltd. of Melbourne, for the valuable time they contributed in training me in the use of their pulsed NMR equipment.

Finally on a more personal note, I must mention the support and encouragement given by my Wife and children who gave up their home and friends in England to make it all possible.

CONTENTS

Page

CHAPTER 4 - RELAXATION PROBES FOR C-TERMINAL SEQUENCE DETERMINATION OF PEPTIDES

CHAPTER 5 - N-TERMINAL SEQUENCING OF PEPTIDES

SUMMARY

Specific magnetic perturbations of ^{1}H and ^{13}C nuclear magnetic resonances caused by the presence of binding paramagnetic ions, have been studied for possible use in the sequence determination of peptides. Perturbing probes may cause dramatic shifts in resonances and/or cause a considerable reduction in relaxation times of nuclei adjacent to the paramagnetic probe. For sequence determinations relaxation probes are generally more useful than shifting probes. Experiments to sequence peptides from the N-terminus using Cu^{2+} and from the C-terminus using Gd^{3+} indicate that it is possible to obtain sequence information of up to five amino acid residues from each site. For C-terminal sequencing, measurement of enhanced spin lattice relaxation is the preferred technique which may also provide some sequence data in cases where resonances overlap.

Studies have been undertaken to improve this sequencing technique in a number of ways which will make it more attractive to the Biochemist. The amount of peptide can be reduced to submilligram quantities by using a specially designed micro NMR tube. It is generally expected that the length over which relaxation probes are effective should increase as the binding constant for a probe at a particular site increases, and therefore, various methods for measuring and improving binding constants of sites in peptides by chemical modification have been examined. In many cases increases in binding constants resulting from modifications appear to be of such magnitude than an insoluble complex precipitates from aqueous solution.

ABBREVIATIONS

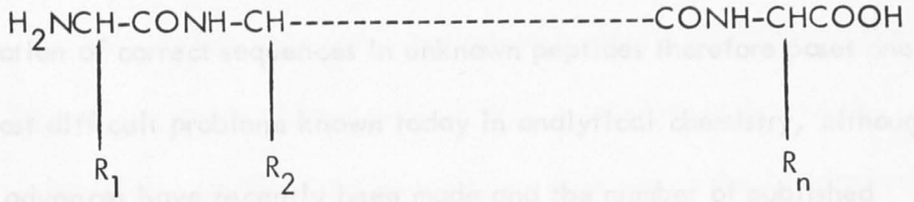
CMR	Carbon-13 nuclear magnetic resonance.
CW	Continuous wave.
di CMB	di-carboxymethyl.
FID	Free induction decay.
FT	Fourier transform.
Gu DCl	Deuterated guanidine hydrochloride.
Gu HCl	Guanidine hydrochloride.
Hz	Hertz (cycles per second).
M	Molar
MHz	Megahertz ($Hz \times 10^6$).
MS	Mass spectrometry.
NMR	Nuclear magnetic resonance.
NOE	Nuclear Overhauser enhancement.
PMR	Proton nuclear magnetic resonance.
ppm	Parts per million.
T_1	Spin lattice relaxation time.
T_2	Spin spin relaxation time.
TMS	Tetramethylsilane.

CHAPTER 1

GENERAL INTRODUCTION

1A. PEPTIDE SEQUENCE DETERMINATION

Peptides are the class of chemical compounds formed from "head to tail" condensation of amino acids, and have the general structure:



where each side chain denoted by R_1, R_2 ----------R_n, varies in structure depending on each particular amino acid in the chain. It has become standard practice to use a three letter shorthand notation for peptides with the amino (N) terminal unit on the left, followed by the complete sequence of amino acids to the carboxyl (C) terminal group on the right. For example, AlaLeuGly would represent the tripeptide alanylleucylglycine with an N-terminal alanine and C-terminal glycine. It should also be noted that each αC in the peptide chain is a chiral atom except in the case of glycine (R=H), but in naturally occurring peptides the L form predominates.

Peptide molecules form the basic building units of protein and enzyme molecules which may consist of many hundreds or thousands of amino acid units all joined together in a precise sequence. It becomes immediately apparent that the number of possible sequences in a specified peptide or protein is very large. For example, a modest decapeptide composed of any of the 20 common amino acids may exist in any of 10^{20} sequences. Evaluation of correct sequences in unknown peptides therefore poses one of the most difficult problems known today in analytical chemistry, although great advances have recently been made and the number of published sequences has increased rapidly over the last few years. The ideal method must be accurate and precise but most important should be adaptable to very small quantities of peptide to satisfy the requirements of biochemists. Existing methods may be broadly classified into chemical and instrumental techniques although there is often no clear distinction between the two because many chemical methods have now been automated, and often instrumental methods require some prior chemical modification of the peptide to alter its physical or chemical properties.

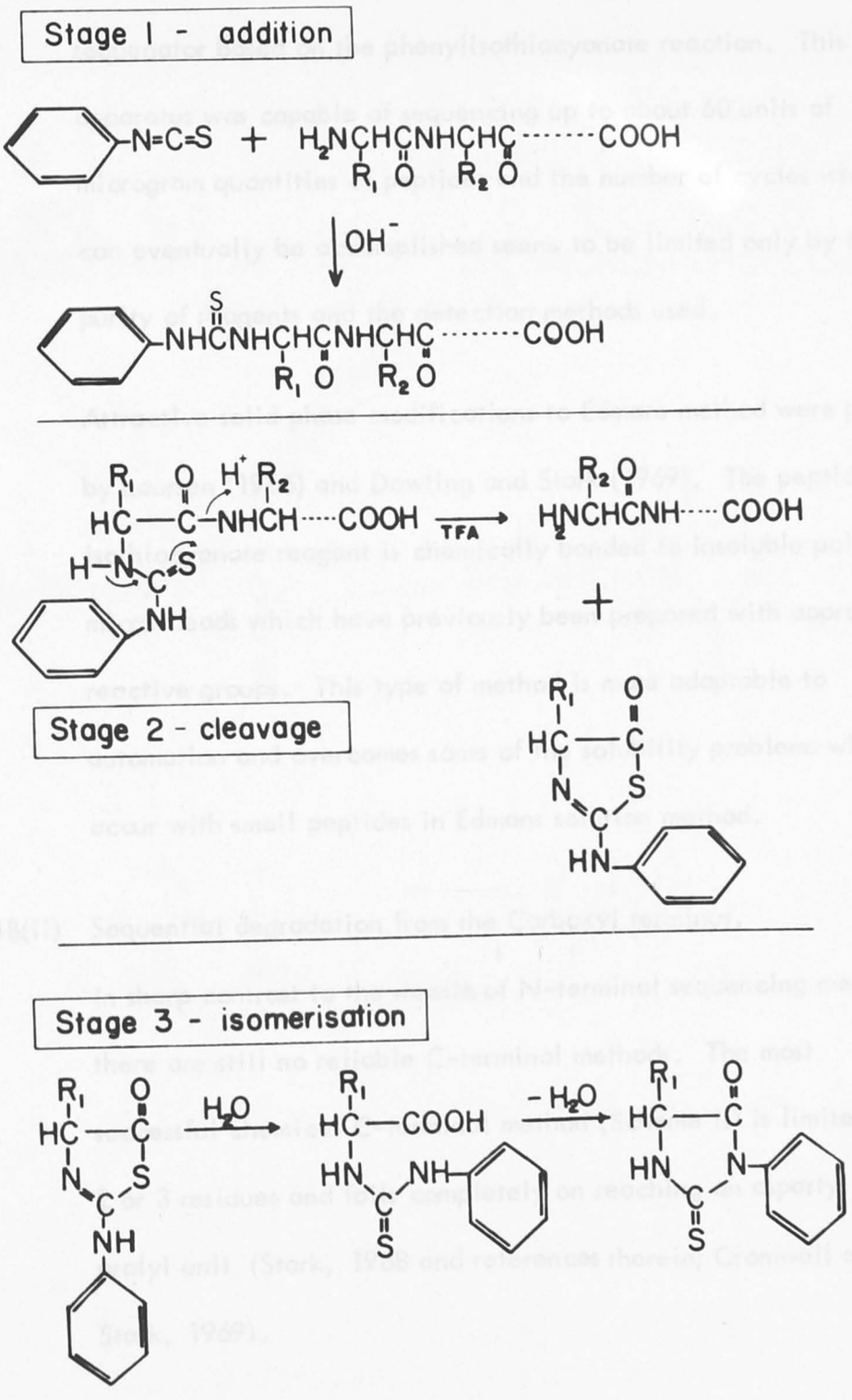
1B. CHEMICAL METHODS

Those which solely identify the N-terminal residue by chemical modification prior to complete hydrolysis are not included here, as they have been adequately reviewed elsewhere (Spande et al, 1976) and are not considered to be complete sequencing methods.

1B(i) Sequential degradation from the N-terminus.

The most important N-terminal chemical method was first proposed by Edman (1950). He evolved a method of stepwise degradation involving reaction of the peptide with phenylisothiocyanate to produce a phenyl thiohydantoin of the N-terminal amino acid (Scheme I). Since the unprotonated α-amino group is the reactive species, initial coupling is performed in mildly alkaline solution. The phenylthiocarbamyl peptide (PTC - peptide) must then be separated from the reaction mixture, after which the N-terminal amino acid residue is cleaved in anhydrous acid and finally isomerised to the more stable phenyl thiohydantoin derivative (PTH). The whole procedure may then be repeated on the remaining peptide. Each PTH-amino acid unit removed must then be identified by either (a) chromatography or electrophoresis of the PTH derivative (b) amino acid analysis of a sample of remaining peptide or (c) chemical identification of the newly formed N-terminal amino acid. All of these methods have found practical use and detailed procedures have been published (Schroeder, 1967; Konigsberg, 1967; Gray, 1967).

More recent studies on the Edman method have concentrated on improving reagents and yields so that longer peptides may be sequenced. Modifications involving the use of other isothiocyanates which yield highly fluorescent hydantoins have proved particularly useful in increasing sensitivity. (Maeda and Kawauchi, 1968).



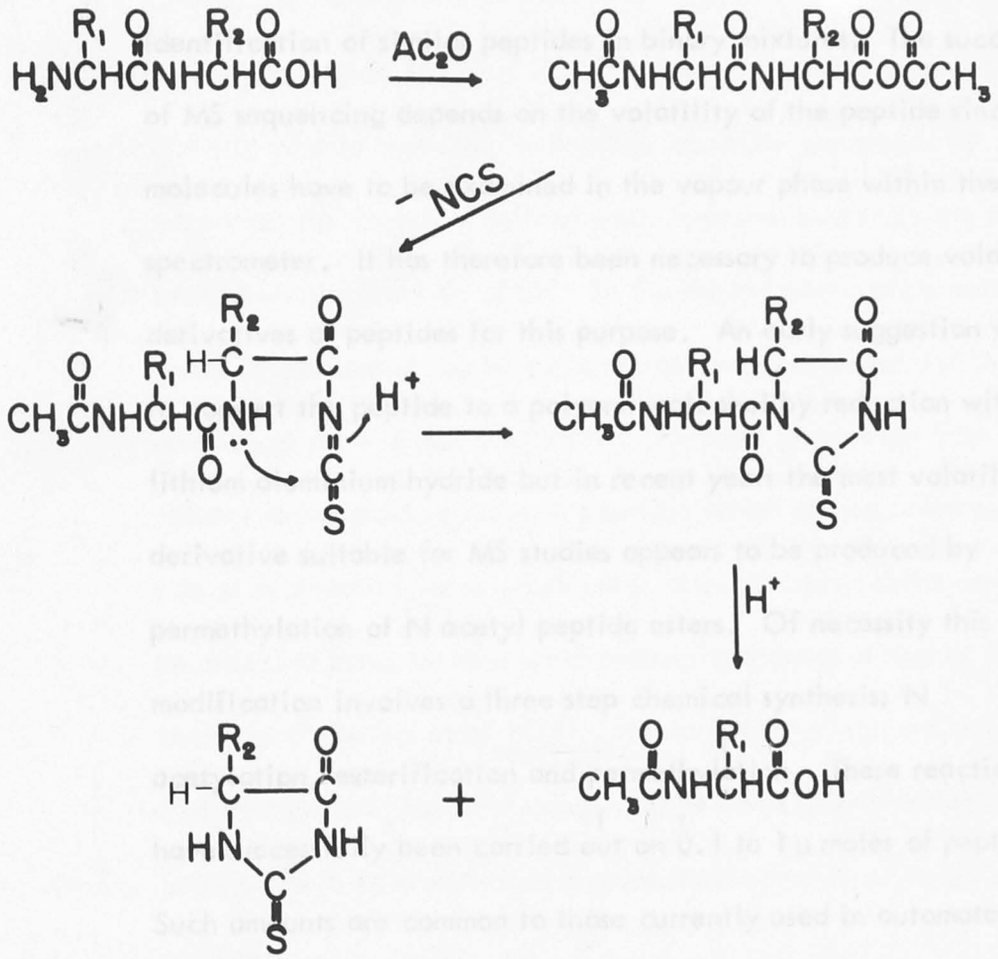
Scheme 1. Edman (1950) method for sequential chemical degradation of peptides from the N-terminus

Efficient sequence determination became a practical reality when Edman and Begg (1967) produced a completely automatic sequenator based on the phenylisothiocyanate reaction. This apparatus was capable of sequencing up to about 60 units of microgram quantities of peptides and the number of cycles which can eventually be accomplished seems to be limited only by the purity of reagents and the detection methods used.

Attractive solid phase modifications to Edmans method were proposed by Laursen (1966) and Dowling and Stark (1969). The peptide or isothiocyanate reagent is chemically bonded to insoluble polymeric micro-beads which have previously been prepared with appropriate reactive groups. This type of method is more adaptable to automation and overcomes some of the solubility problems which occur with small peptides in Edmans solution method.

1B(ii) Sequential degradation from the Carboxyl terminus.

In sharp contrast to the wealth of N-terminal sequencing methods there are still no reliable C-terminal methods. The most successful chemical C-terminal method (Scheme II) is limited to 2 or 3 residues and fails completely on reaching an aspartyl or prolyl unit (Stark, 1968 and references therein; Cromwell and Stark, 1969).



Scheme II. Chemical degradation of peptides from the C-terminus (Stark, 1968).

1C. INSTRUMENTAL METHODS

1C(i) Mass Spectrometry.

Over recent years mass spectrometry has been shown to offer considerable promise for sequence determination of pure oligopeptides, and some claims to success have been made for identification of similar peptides in binary mixtures. The success of MS sequencing depends on the volatility of the peptide since molecules have to be examined in the vapour phase within the spectrometer. It has therefore been necessary to produce volatile derivatives of peptides for this purpose. An early suggestion was to convert the peptide to a polyaminoalcohol by reduction with lithium aluminium hydride but in recent years the most volatile derivative suitable for MS studies appears to be produced by permethylation of N acetyl peptide esters. Of necessity this modification involves a three step chemical synthesis; N acetylation, esterification and permethylation. These reactions have successfully been carried out on 0.1 to 1 μ moles of peptide. Such amounts are common to those currently used in automated chemical methods but because of the lengthy chemical modifications, it seems likely that MS sequencing will only become a method of choice for special cases, such as for an unusual amino acid or an N-terminal blocked amino acid. A major problem with this method is that the permethylation reaction fails in the presence of sulphur containing amino acids, histidine, and arginine, and further chemical modifications are required to these amino acids, prior to the permethylation step.

1C(ii) NMR Spectroscopy

Attempts to obtain information about the primary structure of peptides by NMR spectroscopy are comparatively recent and have concentrated either on changes in resonance shifts with pH or on the use of shift reagents.

Sheinblatt (1966a, 1966b, 1967) pioneered the work on pH studies and was able to show that tripeptides could be sequenced by measuring the chemical shift of each α-proton nuclei in the peptide backbone as a function of pH. In the region where each terminal group is protonated the adjacent nuclei are deshielded and their resonances show a downfield shift. Nuclear resonances from central amino acid residues in peptides which are adjacent to the site of protonation were unaffected, although small shifts can now be detected using modern spectrometers operating at higher field strengths (Anteunis et al, 1974). However this effect reduces so rapidly as a function of distance from the site of protonation that it is unlikely to be useful for sequence determination of longer peptides.

The use of shift reagents for peptide sequencing was studied by Anteunis et al (1973, 1974). Using a combination of pH shifts and aqueous shift reagents at frequencies of 270 MHz to 360 MHz they claim some success in sequencing pentapeptides, but shifts are very small and precipitation of the lanthanide reagent occured at pH values around 7.

A preliminary study on sequential resonance broadening in peptides by Gd^{3+} (Crompton, 1973) appeared useful for peptide sequence determination from the C-terminus and it is on the basis of this work that the present evaluation of NMR relaxation probes has been undertaken. The following theoretical concepts of NMR and the effect of paramagnetic probes are therefore essential.

1D. NMR THEORY

The basic theory of NMR has been adequately covered by many authors (for example see Pople, Schneider, Bernstein, 1959 or Emsley, Feeney, Sutcliffe, 1965). Of more immediate interest is the effect caused by paramagnetic shift or relaxation probes on nuclear magnetic resonances of peptides.

The magnetic resonance of a proton nucleus may be characterised by (a) its chemical shift (ω), (b) its longitudonal relaxation time T_1 and (c) its transverse relaxation time T_2. Each of these values may be affected to a varying extent by the presence of a paramagnetic ion, depending whether it is present in a free or bound state. It is the effect caused by bound paramagnetic ions that are of particular interest for peptide sequence studies because changes in relaxation rates would be expected to vary for different nuclei in the peptide molecule depending on their average distance from the paramagnetic ion.

Effects on Chemical Shifts by paramagnetic ions.

Nuclear resonances from some paramagnetic complexes exhibit large chemical shifts from the positions recorded in similar diamagnetic complexes. This is caused by interaction with unpaired electrons of the paramagnetic ion and may be subdivided into contact and pseudocontact (or dipolar) shifts.

1D(i) Contact shifts

These are often referred to as the Fermi or isotropic contact shifts, and result from the delocalisation of unpaired electron spin density at the resonating nucleus. This effect is transmitted through chemical bonds from the paramagnetic site. Bloembergen (1957) quantitated this phenomenon for the first row transition metals where:

$$\frac{\Delta\omega_m}{\omega_I} = -\left(\frac{A}{\hbar}\right)\frac{g\beta S[S+1]}{3kT\gamma_1}$$

where $\frac{A}{\hbar}$ is the hyperfine coupling constant, g is the Lande g factor, T the absolute temperature, γ_1 the magnetogyric ratio of ^1H nuclei, β is the Bohr magneton, S the total electron spin, $\Delta\omega_m$ is the shift from the position in a corresponding diamagnetic complex and ω_I the irradiating frequency. This equation is only valid for isotropic g values and requires modification for anisotropic g values (Jesson, 1967; Kurland and McGarvey, 1970; Horrocks, 1970).

1D(ii) Pseudo-Contact Shifts

In complexes of isotropic electronic g tensors, the dipolar electron nuclear interaction is averaged out, but for those in which the g-tensor is anisotropic (e.g. paramagnetic lanthanide complexes, except Gd^{3+}), the dipolar interaction results in pseudocontact shifts (McConnell and Robertson, 1958) which for transition metals is given by:

$$\frac{\Delta\omega_m}{\omega_I} = \frac{\beta^2 S[S+1]}{6kT} x r^{-3} F'$$

$$F' = \left\{g_z^2 - g^2\right\}\left\{3\cos^2\theta - 1\right\} + \left\{g_x^2 - g_y^2\right\}\sin^2\theta\cos 2\phi$$

This equation is more frequently encountered in the case of axial symmetry. Because of the many conformations of a molecule in solution resulting from rapid motion on the NMR time scale the values of θ and r result from average of these motions. In this case the above equation is written as:

(McConnell and Robertson, 1958)

$$\frac{\Delta\omega_m}{\omega_I} = D\left\langle \frac{3\cos^2\theta - 1}{r^3} \right\rangle_{av.}$$

The important points to note about the pseudocontact shift are that it is independant of the nucleus but very dependant on the geometry of the complex. The angular dependance predicts shifts of different signs within the same molecule since $(3\cos^2\theta - 1)$ changes signs at an angle of $\theta = 54^\circ 44'$. The through space distance relation ($\propto r^{-3}$) predicts a rapid decrease in shift with distance from the paramagnetic site. In most cases the value of D is unknown and measurement of the ratio of shifts of two nuclei in the same molecule has provided an elegant technique in conformational studies of large molecules (Barry et al, 1971).

1E. RELAXATION IN PARAMAGNETIC SYSTEMS

The presence of paramagnetic ions causes considerable reduction in nuclear relaxation times. This results from the very large magnetic moments of unpaired electrons (approx. 10^3 times that of nuclear magnetic moments). The effects on spin lattice (T_1) and spin spin (T_2) relaxation times by paramagnetic ions have been quantitated by Solomon (1955) and Bloembergen (1957) where :

$$\frac{1}{T_{1M}} = \frac{2}{15}\frac{\gamma_I^2 g^2 S[S+1]\beta^2}{r^6}\left\{\frac{3\tau_C}{1+\omega_I^2\tau_C^2} + \frac{7\tau_C}{1+\omega_S^2\tau_C^2}\right\} + \frac{2}{3}S[S+1]\left(\frac{A}{\hbar}\right)^2\left(\frac{\tau_e}{1+\omega_S^2\tau_e^2}\right)$$

$$\frac{1}{T_{2M}} = \frac{1}{15}\frac{\gamma_I^2 g^2 S\{S+1\}\beta^2}{r^6}\left\{4\tau_C + \frac{3\tau_C}{1+\omega_I^2\tau_C^2} + \frac{13\tau_C}{1+\omega_S^2\tau_C^2}\right\} + \frac{1}{3}S\{S+1\}\left(\frac{A}{\hbar}\right)^2\left(\frac{\tau_e}{1+\omega_S^2\tau_e^2} + \tau_e\right)$$

The first terms of each equation arise from dipolar interactions between electron spin and nuclear spin characterised by a correlation time τ_c, and are approximately equal. The second terms are from the contact (or scalar) interaction characterised by a correlation time τ_e. Hence it can be shown that the ratio of T_1/T_2 for purely dipolar interactions should be 7/6 or 1.17 and values in excess of this would indicate the importance of a scalar interaction.

The dipolar and contact interactions are both time dependant processes which fluctuate with characteristic correlation times of τ_c and τ_e respectively.

where $$\frac{1}{\tau_c} = \frac{1}{\tau_s} + \frac{1}{\tau_M} + \frac{1}{\tau_R}$$

and $$\frac{1}{\tau_e} = \frac{1}{\tau_s} + \frac{1}{\tau_M}$$

where rotation of the complex is characterised by a rotational correlation time (τ_R), electron relaxation by an electron spin correlation time (τ_S) and chemical exchange of ligands between co-ordination sphere and bulk environment by τ_M, an exchange correlation time which is equivalent to the lifetime of a ligand in the co-ordination sphere of the metal ion.

It is important to realise that these correlation times can be temperature and/or frequency dependent where:

(a) τ_M the lifetime of nucleus in the bound state is temperature dependant according to the Eyring relation

$$\frac{1}{\tau_M} = \frac{kT}{h}\left[\exp\left(-\frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R}\right)\right]$$

where k is the Boltzman constant, T the temperature, $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ the enthalpy and entropy of activation.

(b) τ_R is also temperature dependant according to the equation

$$\tau_R = \tau_R^0 \exp\left\{E_R / RT\right\}$$

where E_R is an activation energy for rotational motion.

(c) τ_S shows a much more complicated relationship with both temperature and frequency (Bloembergen and Morgan, 1961) where

$$\frac{1}{\tau_S} = B\left[\frac{\tau_V}{1+\omega_S^2\tau_V^2} + \frac{4\tau_V}{1+4\omega_S^2\tau_V^2}\right]$$

where τ_v is a correlation time, related to the rate at which zero field splitting is modulated by impact of solvent molecules on the aquo complex, and B containing the value of resultant electronic spin and zero field splitting parameters. As τ_v is frequency dependant then $\frac{1}{\tau_S}$ will show a frequency dependance. In addition $\frac{1}{\tau_S}$ is temperature dependant resulting from that of τ_v given by the relation:

$$\tau_V = \tau_V^0 \exp\left\{E_V / RT\right\}$$

where E_v is the activation energy for motion characterised by τ_v.

1E(i) Modifications to the Solomon (1955) and Bloembergen (1957) equations.

The basic Solomon Bloembergen equations contain a number of assumptions that have required many modifications, the more important of which are:

(a) Allowance for more than one electron spin relaxation time. i.e. T_S should ideally consist of spin spin (T_{2S}) and spin lattice (T_{1S}) relaxation times. Modifications have been proposed by Connick and Fiat (1966), and Reuben, Reed and Cohen (1970) to allow for this factor.

(b) For anisotropic rotation (Woessner, 1962) the overall correlation time τ_R may be involved in a function of re-orientation times of different axes.

(c) Anisotropic g values of the electron. Modified equations to cover this point have been proposed by Sternlicht (1965) Karger and Pfeifer (1968).

1F. CHEMICAL EXCHANGE IN PARAMAGNETIC SYSTEMS

Nuclei may exist either in the bulk media or bound to a paramagnetic ion and in the case where chemical exchange occurs between the two sites, relaxation of the bulk media nuclei are affected according to the equations of Swift and Connick, (1962) and Luz and Meiboom (1964):

$$\frac{1}{P_M q T_{1p}} = \frac{1}{P_M q}\left(\frac{1}{T_1} - \frac{1}{T_{1A}}\right) = \frac{1}{P_M q\, T''_{1A}} + \frac{1}{\{\, T_{1M} + \tau_M \,\}}$$

$$\frac{1}{P_M q\, T_{2p}} = \frac{1}{P_M q}\left(\frac{1}{T_2} - \frac{1}{T_{2A}}\right) = \frac{1}{P_M q T''_{2A}} + \frac{1}{\tau_M}\left[\frac{\frac{1}{T_{2M}}\left(\frac{1}{T_{2M}} + \frac{1}{\tau_M}\right) + \Delta\omega^{2}_{M}}{\left(\frac{1}{T_{2M}} + \frac{1}{\tau_M}\right)^2 + \Delta\omega^{2}_{M}}\right]$$

where A and M denote values in unbound and bound sites respectively, $\frac{1}{T''_{1A}}$ and $\frac{1}{T''_{2A}}$ the relaxation rates from dipolar interaction with solvated paramagnetic ion. P_M is the mole fraction of ligand bound to paramagnetic ion in solution and q the co-ordination number. $\Delta\omega_M$ is the chemical shift difference between each environment.

Relaxation rates may be temperature or frequency dependant according to whether chemical exchange is slow, intermediate or fast on the NMR time scale. The temperature dependance arises from the dependances of $\Delta\omega_m$, τ_M, T_{1M} and T_{2M} and that of T_{1M} and T_{2M} on the temperature dependance of τ_M, τ_S, and τ_R. These effects have been discussed in detail (Dwek, 1973) and may be summarised as follows:

Slow exchange region - Where $\Delta\omega^2 > 1/T^2_{2M}$, $1/\tau_M$ and relaxation is governed by rate of chemical exchange of molecules between bulk and bound site. Also $1/\tau_M$ is independant of frequency in this region.

Intermediate exchange - Increase in temperature leads to exchange resonance narrowing. $1/T_{2p}$ is determined by change in precessional frequency where

$$\frac{1}{\tau_m^2} \gg \Delta\omega_m^2 \gg \frac{1}{T_{2m}\tau_m}$$

which gives

$$\frac{1}{P_m q T_{2P}} = \frac{1}{P_m q T''_{2A}} + \tau_m \Delta\omega_m^2 \text{ hence}$$

$1/T_{2p}$ depends on $\Delta\omega_m$ (i.e. on the magnetic field).

Fast exchange - Relaxation rates are the weighted average of the two environments.

$$\frac{1}{T_{2m}\tau_m} \gg 1/T_{2m}, \Delta\omega_m^2$$

in this region $1/T_{1p}$ and $1/T_{2p}$ decreases with increasing temperature.

Hence relaxation rates may decrease with $1/T$ in the slow exchange region or increase in other regions. In addition in the fast exchange region, the temperature variation of $1/T_{1p}$ and $1/T_{2p}$ are determined by $1/T_{1M}$ and $1/T_{2M}$ and τ_c variations are equivalent to temperature changes. These variations in the fast exchange region have been summarised by Dwek, (1973) and can be useful in selecting the region into which experimental data fits (Table 1.1).

Table 1.1

Condition	$\omega_1\tau_c < 1$			$\omega_1\tau_c > 1$		
τ_c	τ_R	τ_m	τ_s^{**}	τ_R	τ_m	τ_s^{**}
$\frac{d\{1/T_{1m}\}}{d\{1/T\}}$	+	+	$\pm$	−	−	$\mp$
$\frac{d\{1/T_{2m}\}}{d\{1/T\}}$	+	+	$\pm$	+	+	$\pm$

** For τ_s the lower sign is applicable if τ_s is frequency dependant.

+ sign indicates an increase in the relevant correlation time for the temperature function noted on the left of the table, and a decrease noted by a - sign.

CHAPTER 2

MATERIALS AND METHODS

2A. MATERIALS

Peptides - All di, tri, and tetrapeptides were obtained from Fox chemicals. Longer peptides including Phe Asp Ala Ser Val, Ser Gly Ala Gly Ala Gly and bradykynin were obtained from Scharwz Mann Laboratories Ltd.

Lanthanide salts were purchased, as the nitrate or chloride hydrates, from Koch Light Laboratories and were not further analysed.

Cu^{2+} solutions were prepared from an analytical reagent grade of $CuCl_2$ (B.D.H.). 6M Guanidine deuterochloride solutions were prepared from Schartz Mann guanidine hydrochloride (specially pure for spectroscopy).

Deuteration was achieved by repeated lyophilisation from D_2O (98%). 98% D_2O was a product of the Australian Atomic Energy Commission.

Iodo acetic acid. This material manufactured by B.D.H. is light and heat sensitive. On storage, the presence of iodine soon becomes apparent. Before use, the iodo acetic acid was dissolved in water and iodine extracted with carbon tetrachloride. Pure iodo acetic acid was then recovered by lyophilisation of the aqueous layer protecting the vessel from light.

Isothiocyanate of phthalic acid. This was prepared by Howell (1976).

2B. METHODS

2B(i) NMR Spectrometers. The spectrometers used in this work were:

(a) JEOL NMH - 100. A 100 MHz continuous wave spectrometer for ^{1}H nuclei, situated in the Chemistry Dept., and operated personally. This instrument used standard 5 mm sample tubes at a normal operating temperature of 26°C unless otherwise stated.

(b) Brüker HX90 pulsed Fourier transform spectrometer at 90 MHz for ^{1}H nuclei and 22.63 MHz for ^{13}C nuclei with proton decoupling. This instrument is situated in the Laboratories of H.B. Selby Pty. Ltd., Ferntree Gully Road, Melbourne, and was operated personally after a training period under the supervision of H.B. Selby Staff. This spectrometer uses an internal D_2O lock with standard 5 mm tubes for protons and 10 mm tubes for ^{13}C spectra. Unless otherwise noted, spectra were accumulated in 4K of computer memory with a 0.4 second recycle time. The accumulated FID was transformed in 4K/4K of memory for both proton and ^{13}C work although noise levels for ^{13}C spectra were normally minimised by exponential multiplication of the FID before transformation.

(c) Brüker HX270 pulsed spectrometer. 270 MHz spectra were obtained on this instrument which is situated in the Chemistry Dept., but is owned and operated as a service by the National NMR Centre. Over the last 12 months, which represents the first year of operation of the service, this instrument has been plagued with continual breakdowns and it seems very doubtful if the optimum resolution and sensitivity of this instrument has ever been achieved. The recent changeover from computer hardware to software control together with increase in memory size to 64K has caused further delays to this service in the last few months with a consequent curtailment of sequencing experiments particularly in T_1 measurements for which no working computer program yet exists. It also seems doubtful policy that operators with little background knowledge of NMR can be adequately trained in a short space of time to provide the required specialised services of the centre.

2B(ii) NMR Spectral data

Chemical shifts are quoted in ppm downfield from an external TMS standard for protons and upfield from internal GuHCl for ^{13}C nuclei. For determination of binding constants with lanthanide shift reagents tetramethylammonium chloride has been used as a non binding internal standard.

Peptide solutions for PMR were normally lyophilised a number of times from D_2O so that the HDO resonance resulting from exchange between D_2O and labile protons in the peptide or guanidine was minimised. If the final lyophilisation was carried out in 100% D_2O and the resulting solid then redissolved in 100% D_2O the remaining HDO resonance was reduced to a level which did not interfere with other resonances.

T_1 measurements were performed using the standard Brüker Instruments automatic program with a $(180^\circ - \tau - 90^\circ)_n$ pulse sequence for both 90 and 270 MHz spectrometers. The pulse lengths for 180° and 90° pulses were optimised before each T_1 experiment by minimising and maximising the FID respectively. This adjustment is particularly important when using 6M GuHCl solutions because the 90° pulse width for peptides in this solution has been found to be 2 or 3 times longer than the value for aqueous solutions. T_1 values have all been calculated using a least squares computer fit to the plot τ vs $\ln (A_\infty - A_\tau)$ where τ is the pulse sequence time delay and A_∞ and A_τ the resonance peak heights with delays of ∞ and τ respectively. The program (Appendix I) also prints a table of the value $\ln (A_\infty - A_\tau)$, and in all cases these points have been plotted manually to check that a straight line relationship exists. This has been the case for all measurements where τ exceeds 40 msec. Attempts to use delay times of less than 40 msec. particularly on the 270 MHz

spectrometer always lead to non linearity in the plots of τ vs $\ln (A_\infty - A_\tau)$. Fortunately the range of T_1 values required are such that delay times of over 40 msec. are perfectly adequate.

2B(iii) Potentiometric titrations. The measurement of equilibrium constants for copper complexes of triglycine were performed using a thermostatted titration cell constructed according to the design of Perrin and Sayce (1966). All precautions suggested by these workers to minimise potential drift during titration were taken. Titrant was added with a calibrated Agla micrometer syringe of 0.5 ml capacity and pH readings recorded on a Radiometer TTTIC pH meter equipped with scale expansion unit capable of recording pH values to 0.005 units. For successful computer fitting it was necessary to record some 100 to 150 points over each titration curve.

2B(iv) pH Measurements. Apart from those noted above, all pH measurements were obtained on a Beckman research pH meter using a saturated calomel reference electrode and a glass pH electrode. In the case of measurements in D_2O solution, the pH is recorded as a pH meter reading with no correction for deuterium isotape effect, although the pD can generally be obtained by adding 0.44 to the pH reading (Mikkelsen and Nielsen, 1960) or 0.40 (Bradbury and Scheraga, 1966).

2B(v) Viscosity measurements.

Solution viscosity studies were undertaken using a suspended level Ubbelohde dilution viscometer with a capillary diameter of 0.6m.m. All normal precautions necessary for accurate and precise flow time measurement were taken (Bradbury, 1970). Flow times in this type of viscometer are not dependant on solution volume and hence all series dilutions were carried out by addition of solvent to the viscometer tube. The measurements were made in a constant temperature bath at 25°C and basic precautions were taken, such as vertical alignment of viscometer and filtration of sample solutions before use. Flow time measurements for each solution were repeated until agreement between successive readings of 0.1 secs. was obtained.

2B(vi) Mass Spectrometry

The mass spectra obtained on permethylated peptides were recorded on a Varian CH7 mass spectrometer which is run as a service in the Department of Chemistry.

2B(vii) Computers

All computer programs which have been written for data fitting with iterative procedures or calculations are listed together with explanations in Appendix I. The logic of programs used for binding constants is shown under the appropriate method in Chapter 7. Two computers and associated languages have been used:

The John Curtin School PDP11 with remote terminal situated in the Chemistry Department. This computer uses the simple language FOCAL which is ideal for many repetitive type of calculations but is severly limited on core size, and is unable to handle some of the larger programs.

The A.N.U. Computer Centre UNIVAC 1108. This computer has been used with FORTRAN IV. Data imput was on cards for large programs and on time shared demand terminals for smaller programs.

CHAPTER 3

PARAMAGNETIC IONS AS PROBES FOR THE SEQUENCE DETERMINATION OF PEPTIDES

3A. INTRODUCTION

Probes are usually small molecules or ions which can be chemically or physically attached at unique sites to molecules under study, and are then able to provide structural information about the molecule at or near the binding site. The two types of probe in general use are subdivided into (1) detecting and (2) perturbing probes, although some probes may act as both. Perturbing probes cause a change in some property of the molecule to which it is attached whilst detecting probes show some change in their own properties.

There are many types of magnetic perturbing probes which will affect the resonance position or relaxation rate of a nucleus. For peptide sequence determination some limitations on the type of probes which can be used are immediately apparent (e.g. they must be soluble in the polar solvents normally required for peptides). Additionally those probes which exhibit a pseudocontact effect (distance related) would be preferable, because the terminal carboxyl group in a peptide chain is an ideal binding site, and each peptide unit occurs at an increasing distance from this terminal group when the peptide is present in a random coil configuration. However, other probes which exert a predominantly contact interaction (such as transition metal ions) have not been dismissed at this stage.

The ionic salts of rare earths have long attracted the attention of NMR spectroscopists for use as aqueous shift probes (Conger et al, 1952 Shulman et al, 1959, Reuben et al, 1967, Lewis et al, 1962) because of their short electronic relaxation times and resultant minimal line broadening effects. More recently the success of lanthanide β diketonates (Hinkley et al 1969) as shift reagents in non-aqueous solvents has revitalised interest in lanthanide salts as aqueous probes. Induced shifts as large as 4 ppm have been recorded using lanthanide ions bound to carboxylic acids (Hart et al, 1971), amino acids (Robb et al, 1972), steroids (Herts et al, 1971), nitrones (Sanders et al, 1972), phosphates (Sanders et al, 1972), and esters (Reyes-Zamora et al, 1971). In fact the popularity of NMR shift reagents for many diverse applications is now so great that many lengthy review articles have recently been published (Reuben, 1973; Nieboer, 1975; Cockerill, 1973).

The lanthanides are the series of elements with atomic numbers between 58 and 71 in which the 4f electronic shell is progressively filled. The 4f electrons are effectively shielded by the outer 5s and 5p orbitals and complex formation is therefore largely due to electrostatic attraction with little covalency.

Although the lanthanides are of little biological importance their ionic radii ranging from 1.061Å for La to 0.848Å for Lu is well within the range of the ionic radius of Ca^{2+} (0.99Å) and the usefulness of lanthanides as spectroscopic and magnetic probes in biological systems after isomorphous replacement is now well recognised (Williams, 1970).

3A(i) Classification of paramagnetic probes.

In particular, paramagnetic ions may produce resonance shifts or cause enhanced relaxation of adjacent nuclei to the site of binding of the ion, (i.e. reduction in T_1 and T_2). Changes in T_1 may be directly measured by a number of pulse sequence techniques (Vold et al, 1968, Freeman et al, 1970, 1971, Markly et al, 1971) whilst changes in T_2 may be conveniently monitored by changes in resonance line width ($\Delta\nu$) in Hz, using the relation $1/T_2 = \pi \Delta \nu$. All paramagnetic ions give rise to both shifts and changes in relaxation but the magnitude of either effect depends on τ_c . The following empirical classification of paramagnetic ions into "shifters" and "broadeners" has been proposed:

(1) Ions in which $1/\tau_c$ is dominated by $1/\tau_r$ with $\tau_r \simeq 10^{-10}$ secs. This group of ions cause changes in relaxation (broadening) and include Mn(II), Gd(III), Eu(II), V(II), Cu(II).

(2) Ions where $1/\tau_c$ is dominated by $1/\tau_s$ with $\tau_s \simeq 10^{-12}$ secs. This group produces shifts and include Co(II), Fe(II), Fe(III), Ni(II), and most lanthanide (III) ions excluding Gd(III).

Group (1) ions have $1/T_{1M}$ values at least two orders of magnitude greater than group (2) and hence their usefulness as relaxation probes.

For peptide sequence determination there is the choice between shifting or broadening probes, and the ensuing comparison has been undertaken to assess the differences, advantages and limitations of each group.

3B. SHIFTING PROBES

In view of the predominantly ionic character of lanthanide complexes, they should be useful as shifting probes for peptides by binding to the ionised C-terminal site. This will require working at a pH above the pK of the carboxyl group (>3) and below the pH at which the lanthanide will precipitate as its hydroxide. Ions such as Pr^{3+}, Nd^{3+}, Td^{3+}, Dy^{3+} and Ho^{3+} cause downfield shifts whilst Eu^{3+}, Er^{3+}, Tm^{3+} and Yb^{3+} cause upfield shifts. The most useful ions are those which cause shifts without broadening (Eu^{3+} and Pr^{3+}). Other ions may produce a greater shift but the concomitant line broadening makes spectral assignment difficult.

In general the α-proton resonances of peptides in a denaturant (e.g. 6M GuDCl) occur over a 1 ppm range upfield from the HDO resonance with β side chain resonances still further upfield. Shift reagents which produce downfield shifts of the α-proton resonances of peptides have been found more useful, because they cause less spectral overlap between shifted α with β resonances, although at certain concentrations of shift reagent the resonance can be "lost" under the HDO resonance.

In shifting experiments particularly where K values are required it is necessary to record the resonance shift as a function of added reagent concentration which although simple in principle, is often difficult in practice. Shifts must be measured from an internal non binding standard because changes in magnetic susceptibility prohibit the use of an external standard. Such internal standards are sometimes difficult to find for certain solvents. Quite often it is necessary to carry out very many additions of shift reagent so that resonances may be correctly assigned during "crossover". This latter point is illustrated in fig. 3.1 where in the comparatively simple spectra of tetraglycine (+ Pr^{3+}) it is difficult to follow the C-terminal α-proton resonance as it traverses the other resonances. It should also be noted that only the C-terminal and adjacent α-proton resonance of tetraglycine show any measurable shift although it has been claimed possible (Anteunis M. et al, 1973) to detect shifts for all α-proton resonances of some tetrapeptides using high field spectrometers. The high molar ratios of Pr^{3+} to peptide required to achieve maximum shift (fig. 3.1) result from the low binding constant, and even at 9 times molar excess the shift of the fully bound form of tetraglycine has not been achieved.

With non-glycine peptides where spin spin spitting of the α-proton resonances are observed the shifted spectra are very complex. The complicated resonance patterns resulting from overlap as shifted multiplets cross one another often result in uninterpretable spectra. At large concentrations of reagent some simplification is observed, but unless the shifts can be followed initially at low reagent concentrations then incorrect assignments can often result.

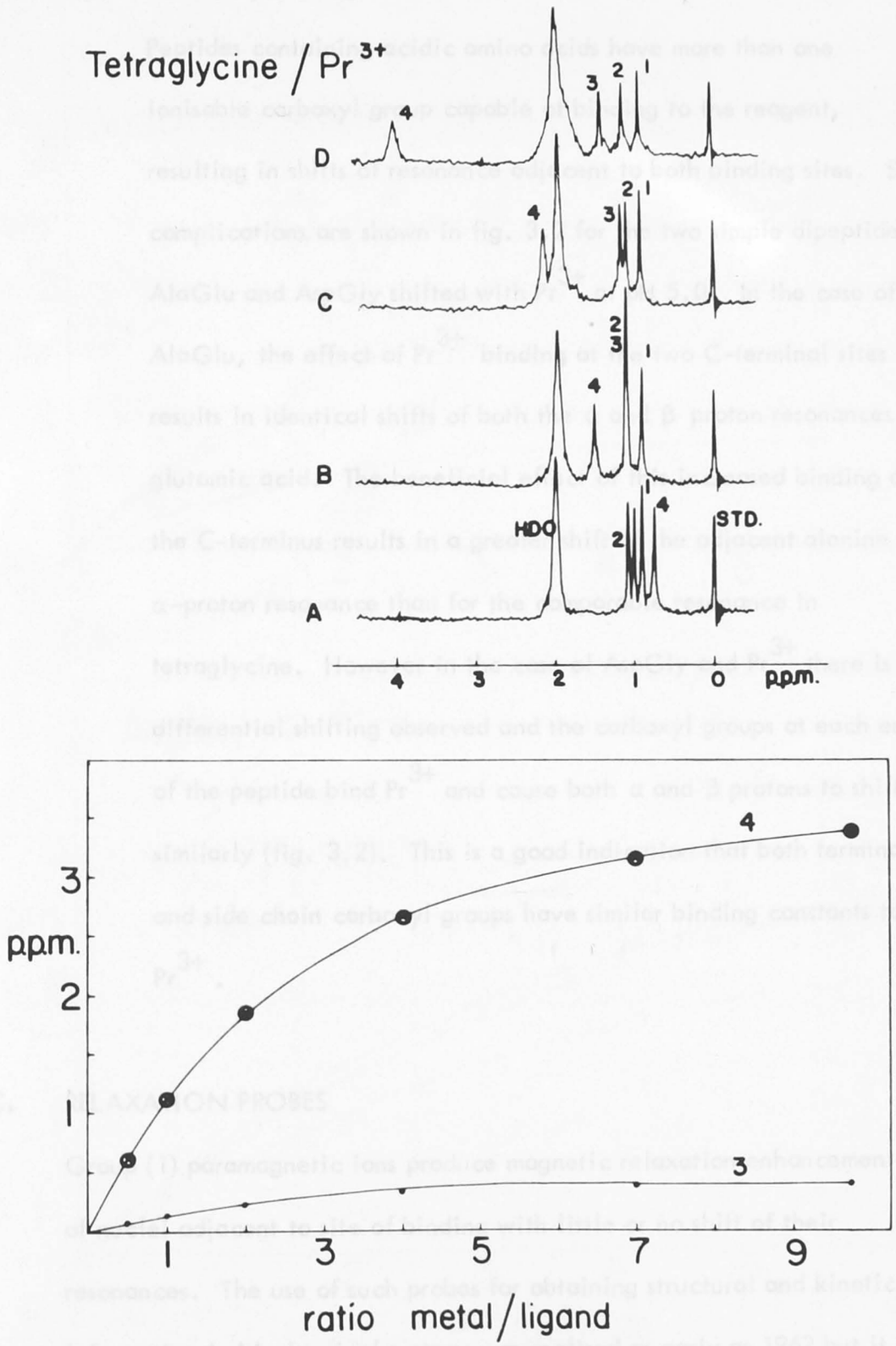


Fig. 3.1 Top, 100 MHz NMR spectra of the α-proton region of tetraglycine (0.1M) with addition of Pr^{3+}, at pH 5.0 in D_2O. Shifts are quoted in ppm from internal standard (tetramethylammonium chloride). Molar Ratio of Pr^{3+}/tetraglycine (A) No. Pr^{3+}, (B) 1.0, (C) 1.5, (D) 10.0 Resonances are numbered from the N-terminus.

Bottom, Plot of shifts for resonances 4 and 3 against ratio of Pr^{3+}/ tetraglycine. No shift was measurable for resonance 1 and 2.

3B(i) Secondary binding sites.

Peptides containing acidic amino acids have more than one ionisable carboxyl group capable of binding to the reagent, resulting in shifts of resonance adjacent to both binding sites. Such complications are shown in fig. 3.2 for the two simple dipeptides AlaGlu and AspGly shifted with Pr^{3+} at pH 5.0. In the case of AlaGlu, the effect of Pr^{3+} binding at the two C-terminal sites results in identical shifts of both the α and β proton resonances of glutamic acid. The beneficial effect of this increased binding at the C-terminus results in a greater shift of the adjacent alanine α-proton resonance than for the comparable resonance in tetraglycine. However in the case of AspGly and Pr^{3+} there is no differential shifting observed and the carboxyl groups at each end of the peptide bind Pr^{3+} and cause both α and β protons to shift similarly (fig. 3.2). This is a good indication that both terminal and side chain carboxyl groups have similar binding constants to Pr^{3+}.

3C. RELAXATION PROBES

Group (1) paramagnetic ions produce magnetic relaxation enhancement of nuclei adjacent to site of binding with little or no shift of their resonances. The use of such probes for obtaining structural and kinetic information in biochemical systems was realised as early as 1962 but it is only recently that Gd^{3+} has become very popular for structural studies. The usefulness of Gd^{3+} as a pseudocontact broadening probe for the sequence

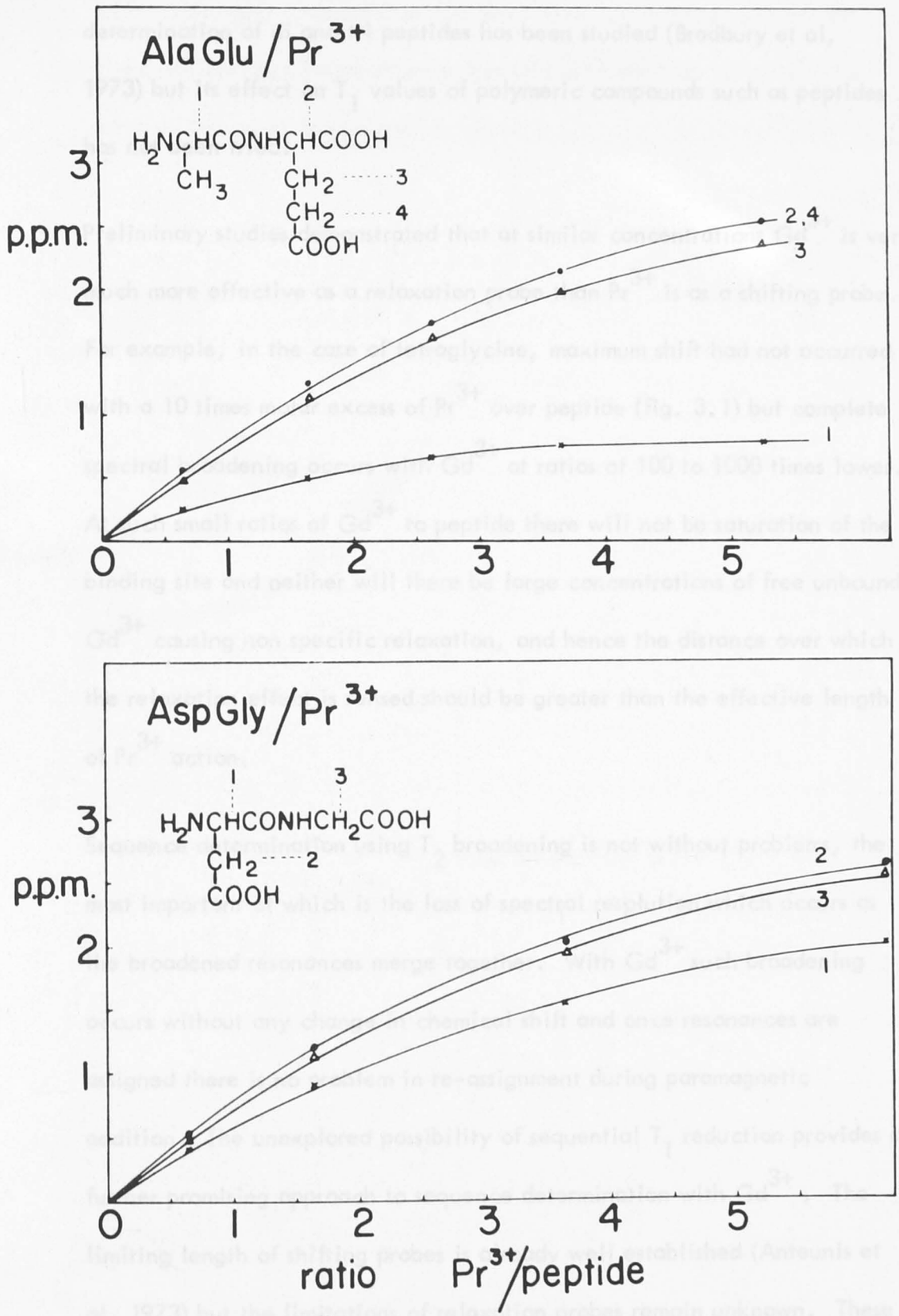


Fig. 3.2 Shifts of proton resonances of AlaGlu (0.1M) and AspGly (0.1M) upon addition of Pr^{3+} at pH 5.0. The numbers on each graph correspond to the numbered nuclei shown in the structures of each dipeptide.

determination of di and tri peptides has been studied (Bradbury et al, 1973) but its effect on T_1 values of polymeric compounds such as peptides has not been tried.

Preliminary studies demonstrated that at similar concentrations Gd^{3+} is very much more effective as a relaxation probe than Pr^{3+} is as a shifting probe. For example, in the case of tetraglycine, maximum shift had not occurred with a 10 times molar excess of Pr^{3+} over peptide (fig. 3.1) but complete spectral broadening occurs with Gd^{3+} at ratios of 100 to 1000 times lower. At such small ratios of Gd^{3+} to peptide there will not be saturation of the binding site and neither will there be large concentrations of free unbound Gd^{3+} causing non specific relaxation, and hence the distance over which the relaxation effect is sensed should be greater than the effective length of Pr^{3+} action.

Sequence determination using T_2 broadening is not without problems, the most important of which is the loss of spectral resolution which occurs as the broadened resonances merge together. With Gd^{3+} such broadening occurs without any change in chemical shift and once resonances are assigned there is no problem in re-assignment during paramagnetic addition. The unexplored possibility of sequential T_1 reduction provides a further promising approach to sequence determination with Gd^{3+}. The limiting length of shifting probes is already well established (Anteunis et al, 1973) but the limitations of relaxation probes remain unknown. These factors combine to make relaxation probes more attractive than shifting probes for peptide sequence determination.

In the following chapters attempts have been made to answer some of the questions about relaxation probes and to assess their suitability for peptide sequence determination.

CHAPTER 4

RELAXATION PROBES FOR C-TERMINAL SEQUENCE DETERMINATION OF PEPTIDES

4A. INTRODUCTION

It is well known that the paramagnetic ground state ions, Gd^{3+} and Eu^{2+}, possess long electron spin relaxation times compared to those of other rare earth ions and to their own rotational correlation time (τ_R). This is a basic requirement for NMR relaxation probes. Paramagnetic contributions to proton relaxation should according to the Solomon (1955) and Bloembergen (1957) equations show a distance related effect which is proportional to r^{-6} where r is the distance of the observed nuclei from the paramagnetic centre. Where compounds are able to bind lanthanide ions at a specific site and undergo rapid and reversible exchange with the bulk media then a progressive decrease in T_1 and T_2 with increasing concentration of lanthanide should result. Rapid exchange on the NMR time scale provides a mechanism for transfer of the effects of Gd^{3+} over all ligands even when the Gd^{3+} ion is present at very low concentration compared to that of the ligand. In such cases the proton resonances actually observed in the NMR spectra are those of the unbound ligands whilst those of bound ligands are too broad for observation.

4B. APPLICATION TO PEPTIDE SEQUENCING

In peptide molecules the C-terminal carboxyl group in its ionised form provides a unique binding site for lanthanide ions. For application of the distance related NMR paramagnetic relaxation effect to peptide sequence determination each amino acid unit must occur in solution at an increasing distance from this terminal binding site. In the case of peptides this condition is fulfilled when the molecule is in a random coil configuration which may be achieved by using a denaturing solution such as 6M GuHCl or 8M urea. In this configuration the distance of each amino acid unit from the terminal group increases in proportion to $\sqrt{n}$ where n is the residue number from the C-terminus, hence the paramagnetic relaxation term might then be expected to show an r^{-3} relation with number of peptide units.

A serious difficulty encountered with relaxation probes such as Gd^{3+} is their ability to cause non specific relaxation of all nuclei, when present as free unbound ions. To achieve maximum length effectiveness in sequencing studies it is necessary to minimise this general broadening effect by adjusting solution conditions where at all possible to maximise binding. The formation constants for binding of Gd^{3+} to carboxyl groups are between 1 and 50. Assuming 1 : 1 complex formation it is possible to calculate the amount of free and bound Gd^{3+} for various binding constants and for varying ligand concentrations.

Such calculations show (fig. 4.1) that maximum binding is achieved by:

(1) Keeping the total ligand concentration as high as possible.

(2) Using binding sites on the ligand with the highest possible formation constants.

Furthermore binding will also be improved by maintaining a pH well above the pK of the carboxyl group. Attempts to incorporate improved binding sites in peptides will be discussed in a later chapter.

4B(i) Relaxation mechanism.

A predominantly dipolar mechanism is essential if one is to use the distance related paramagnetic relaxation effect. A convenient way of establishing the presence of a dipolar interaction is to measure the ratio of T_1 / T_2 in the presence of increasing concentrations of paramagnetic ion (Espersen et al, 1976). In the model system of triglycine/Gd^{3+} the ratios of T_1 / T_2 have been measured and range from 1.96 to 1.50 which is close to the theoretical value of 1.17 for a purely dipolar interaction.

4B(ii) NMR of peptides in a denaturing solvent (6M GuDCl).

All sequence experiments require the use of 6M GuDCl/D_2O to maintain a random coil configuration of the peptide in solution. The pH of solutions was 5.0 which was high enough to ensure deprotonation of the terminal carboxyl group but low enough to prevent precipitation of $Gd(OH)_3$. An advantage of this solvent is the increased solubility exhibited by a number of peptides which are only sparingly soluble in D_2O alone. There are, however, some

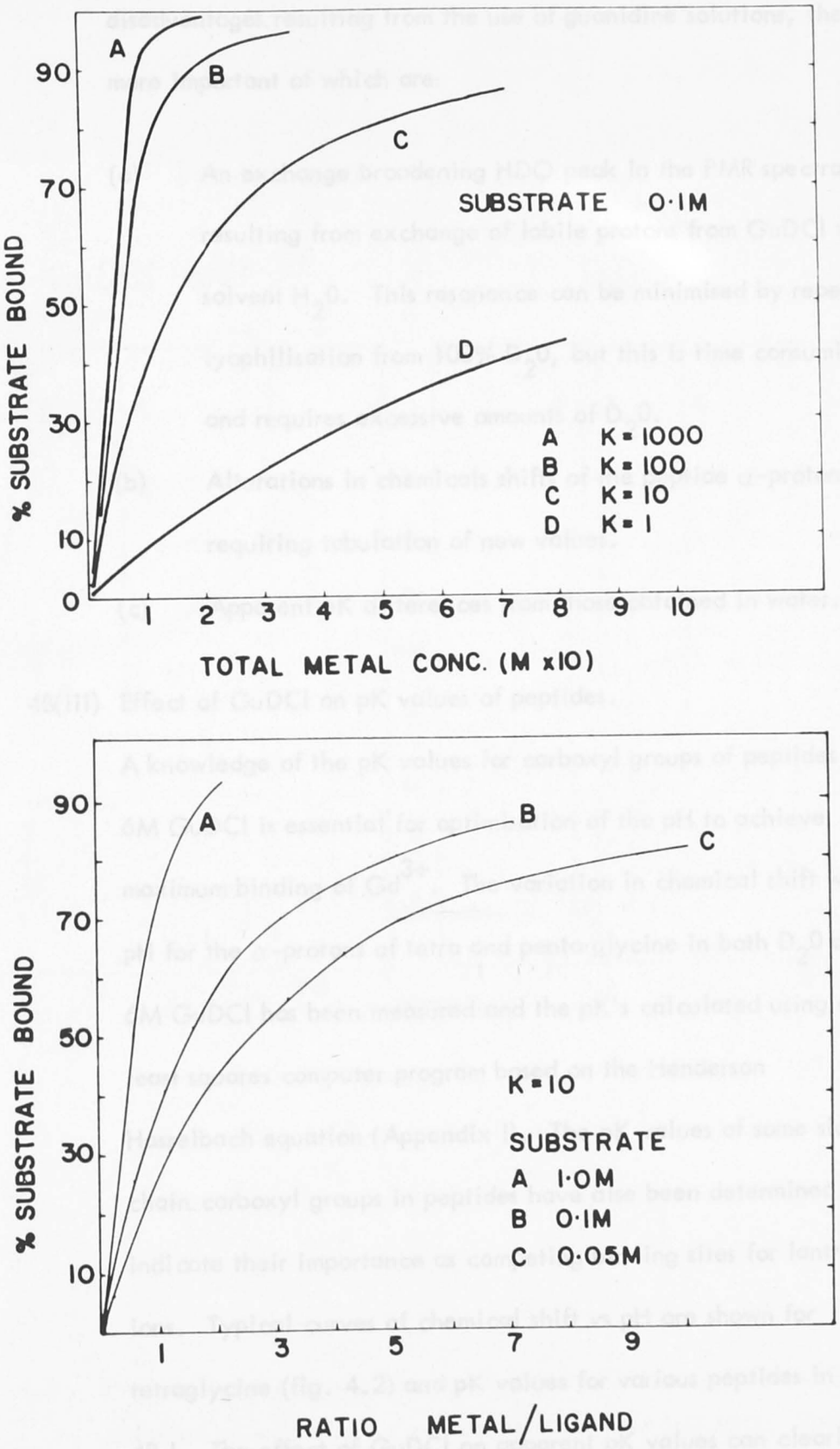


Fig. 4.1 Theoretical binding curves for 1 : 1 complexes.

Top: Variation of binding constants with fixed substrate and and lanthanide concentration.

Bottom: Variation in substrate concentration for fixed binding constant (10) and lanthanide concentration.

disadvantages resulting from the use of guanidine solutions, the more important of which are:

(a) An exchange broadening HDO peak in the PMR spectra resulting from exchange of labile protons from GuDCl with solvent H_2O. This resonance can be minimised by repeated lyophilisation from 100% D_2O, but this is time consuming and requires excessive amounts of D_2O.

(b) Alterations in chemicals shifts of the peptide α-protons requiring tabulation of new values.

(c) Apparent pK differences from those obtained in water.

4B(iii) Effect of GuDCl on pK values of peptides.

A knowledge of the pK values for carboxyl groups of peptides in 6M GuDCl is essential for optimisation of the pH to achieve maximum binding of Gd^{3+}. The variation in chemical shift with pH for the α-protons of tetra and penta glycine in both D_2O and 6M GuDCl has been measured and the pK's calculated using a least squares computer program based on the Henderson Hasselbach equation (Appendix I). The pK values of some side chain carboxyl groups in peptides have alse been determined to indicate their importance as competing binding sites for lanthanide ions. Typical curves of chemical shift vs pH are shown for tetraglycine (fig. 4.2) and pK values for various peptides in table 4B I. The effect of GuDCl on apparent pK values can clearly be seen. The values of side chain and terminal carboxyl groups tend

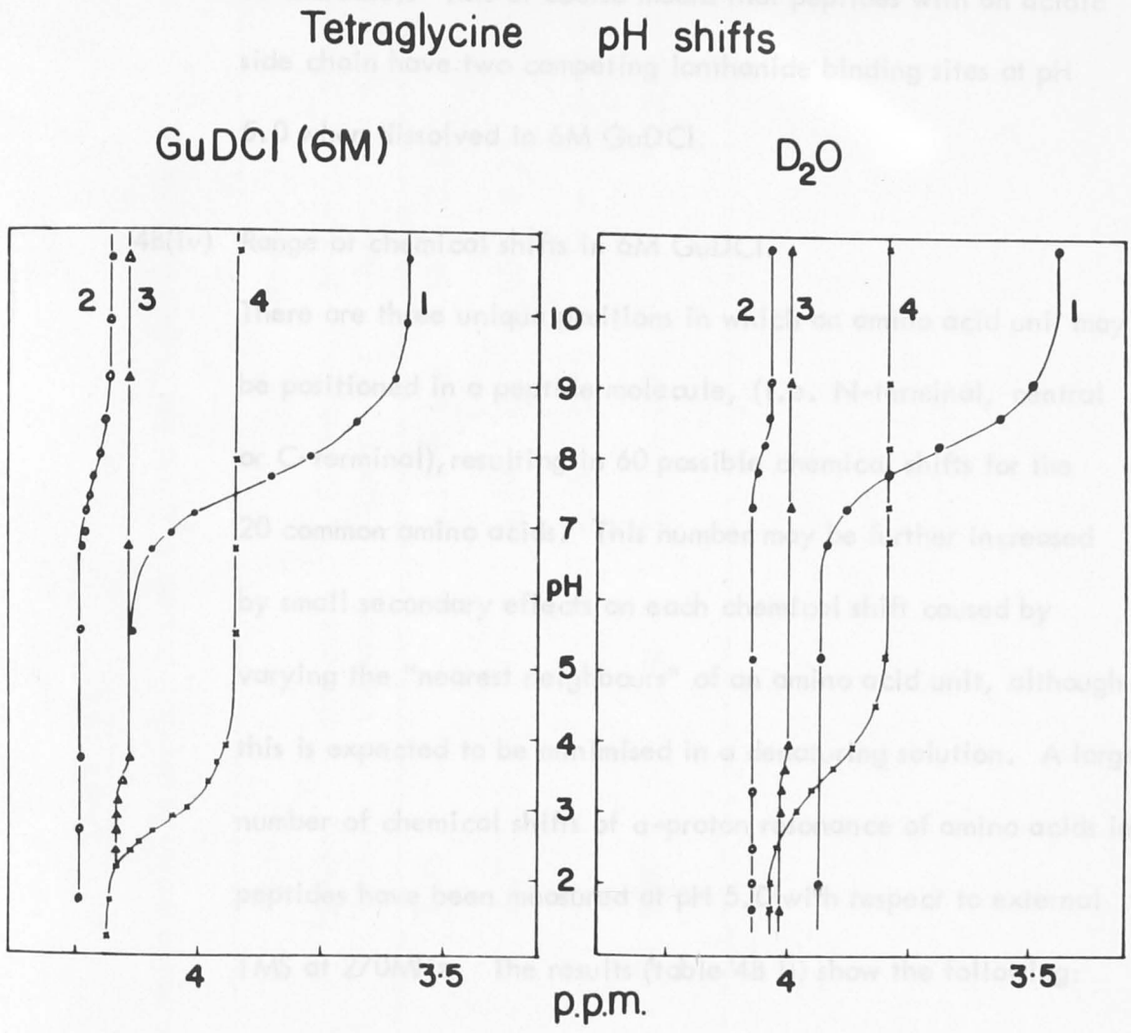


Fig. 4.2 PMR shifts for tetraglycine as a function of pH in GuDCl and D_2O obtained on a continuous wave spectrometer at 100 MHz. Shifts are quoted in ppm against an external TMS standard. Numbers on each curve refer to the α-protons of tetraglycine numbered sequentially from the N-terminus.

towards a common value (side chain carboxyl pK values show a decrease compared to pure D_2O, and terminal carboxyl pK values an increase). This of course means that peptides with an acidic side chain have two competing lanthanide binding sites at pH 5.0 when dissolved in 6M GuDCl.

4B(iv) Range of chemical shifts in 6M GuDCl.

There are three unique positions in which an amino acid unit may be positioned in a peptide molecule, (i.e. N-terminal, central or C-terminal), resulting in 60 possible chemical shifts for the 20 common amino acids. This number may be further increased by small secondary effects on each chemical shift caused by varying the "nearest neighbours" of an amino acid unit, although this is expected to be minimised in a denaturing solution. A large number of chemical shifts of α-proton resonance of amino acids in peptides have been measured at pH 5.0 with respect to external TMS at 270MHz. The results (table 4B II) show the following:

(a) An overall range about 1 ppm for the α-proton resonance which is less than the 2 ppm range found in D_2O alone.

(b) Considerable overlap of α-proton resonances from different amino acids. Even with 270 MHz spectrometers it is very unlikely that all resonances will be resolved.

(c) The "nearest neighbour effect" is of the order of 0.2 ppm, determined from glycine containing peptides (Table 4B III).

TABLE 4B I

pK values for carboxyl and amino groups in peptides in both D_2O and 6M GuDCl solutions

	D_2O			6M GuDCl		
	C-term. Carboxyl	N-term. Amino	Side chain Carboxyl	C-term. Carboxyl	N-term. Amino	Side chain Carboxyl
Tetraglycine	3.05	7.75		3.14 $\pm$ 0.2	7.68 $\pm$ 0.17	
Pentaglycine	3.05	7.70		3.10 $\pm$ 0.2	7.50 $\pm$ 0.2	
PheAspAlaSerVal	*	*	*	3.30 $\pm$ 0.15	*	3.48 $\pm$ 0.1
AspGly	2.10	9.07	4.45	3.50 $\pm$ 0.11	*	3.50 $\pm$ 0.2

* Values not determined

TABLE 4B II

Chemical shifts of α ^{1}H resonances for peptides dissolved in 6M GuDCl at pH 5.0.

Shifts were measured against external TMS at 270 MHz.

Amino acid	Peptide	p.p.m.
Tyrosine	TyrGlyGly	3.37
	GlyTyrGly	3.74
	ValTyrVal	3.77
	GluTyrGlu	3.77
	GlyGlyTyr	3.29
	ThrLysTyr	3.39
Valine	ValTyrVal	3.86
	ValTyrVal	3.89
	PheAspAlaSerVal	3.88
Alanine	AlaLeuGly	3.39
	AlaAlaAla	3.19
	AlaAlaAla	3.38
	GlyAlaPhe	3.34
	AlaAlaAla	3.03
	GlySerAla	3.30
	MetMetAla	3.07
Leucine	LeuGlyPhe	3.12
	AlaLeuGly	3.25
	GlyLeuGlyLeu	3.30
	GlyLeuGlyLeu	3.15

TABLE 4B II

(continued)

Amino acid	Peptide	p.p.m.
Phenylalanine	PheGlyGly	3.41
	GluGlyPhe	3.35
	GlyAlaPhe	3.34
	LeuGlyPhe	3.37
Arginine	ArgProArg	2.70
	ArgProArg	2.98
	GlyGlyArg	3.09
Isoleucine	IleIleIle	3.11
	IleIleIle	3.39
	IleIleIle	2.99
Lysine	LysTrpLys	2.90
	ThrLysTyr	2.92
	LysTrpLys	3.03
Proline	ArgProArg	3.35
	GlyGlyPro	3.27
Threonine	ThrGlyGly	3.28
	ThrThrThr	3.28
	ThrThrThr	3.59
	ThrThrThr	3.26

TABLE 4B II

(continued)

Amino acid	Peptide	p.p.m.
Glutamic acid	GluGlyPhe	3.15
	GluGluGlu	3.02
	GluGluGlu	3.40
	GluGluGlu	3.16
	GlyGlyGlu	3.14
Serine	SerGlyGly	3.33
	GlySerAla	3.71
	GlyAspSer	3.38
	GlyGlySer	3.25
Methionine	MetMetAla	3.52
	MetMetAla	3.52
	GlyGlyMet	3.23
Tryptophan	TrpGlyGly	3.50
	LsyTrpLys	3.72
	GlyTrpGly	3.71

TABLE 4B III

Chemical shifts of α ^{1}H resonances of glycine residues in peptides at pH 5.0 in 6M GuDCl at 270 MHz against external TMS

Peptide	Shift of α-proton resonance (p.p.m.)		
	N-terminal	Central	C-terminal
GlyAlaPhe	2.88	-	-
GlySerAla	3.13	-	-
GlyGlyMet	2.96	3.05	-
GlyGlyGlu	2.98	3.06	-
GlyGlySer	2.97	3.08	-
GlyGlyTyr	3.06	3.10	-
GlyGlyArg	2.95	3.05	-
GlyGlyPro	3.02	3.20	-
GlyGlyPhe		2.99	-
LeuGlyPhe	-	3.01	-
PheGlyGly	-	2.96	2.74
SerGlyGly	-	3.08	2.78
AlaLeuGly	-	-	2.77

4C. SEQUENCE DETERMINATION BY RESONANCE BROADENING (REDUCTION IN T_2)

The binding of Gd^{3+} to the C-terminus of a denatured peptide was found to be useful for exerting a sequential reduction in T_2 for protons along the backbone of the peptide chain. The result of this process is a gradual broadening of the α-proton resonances as the Gd^{3+} concentration is increased, which can be conveniently followed by plotting peak heights against added Gd^{3+} concentration and allows prediction of sequence for tripeptides and some tetrapeptides. The difficulty of sequence determination increases as the number of non glycine units decrease due to the spectral complexity, as the resonances change from singlets through to quartets. For example tetraglycine may easily be sequenced by this method (fig. 4.3) whilst broadening experiments with GlyLeuGlyLeu (Bradbury et al, 1973) require a careful analysis of the spectra in order to predict the correct sequence. With the latter peptide the phenomenon of chemical exchange spin decoupling (Frankel, 1969) has been noted. This effect causes collapse of multiplet resonances at the same time as they undergo paramagnetic broadening and it is possible for these two effects to cancel out one another and lead to a non sequential broadening curve. If broadband decoupling could be achieved of all α-protons from all β-protons it should be possible to partially overcome this problem but so far it has not been possible to modify existing equipment to undertake such an experiment. Therefore sequence determination by T_2 broadening can only be considered a suitable method for very simple peptides preferably

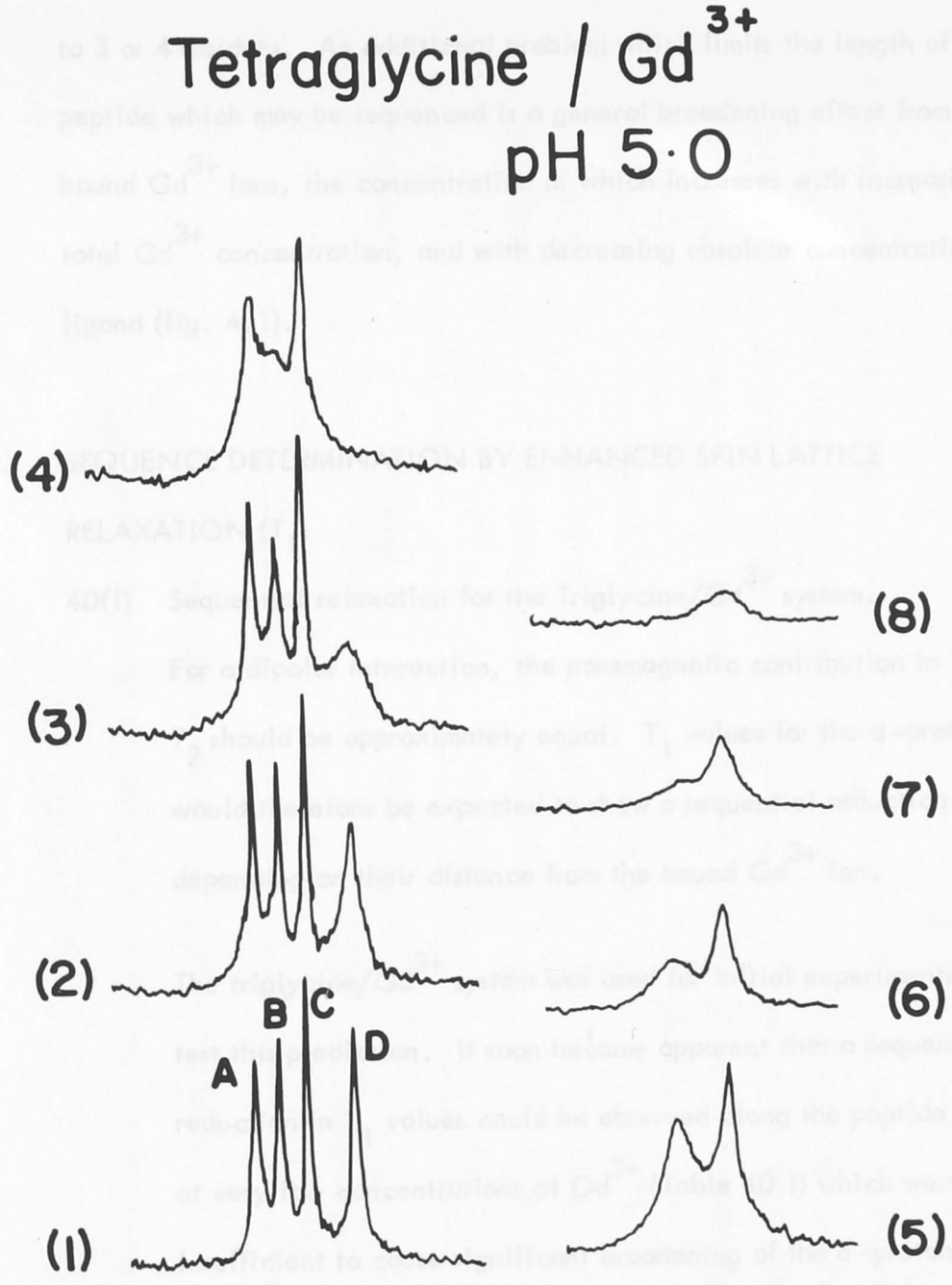


Fig. 4.3 Continuous wave ^{1}H NMR spectra of the α-proton region of tetraglycine with increasing concentrations of Gd^{3+} of (1) NIL (2) 1×10^{-5}M (3) 5×10^{-5}M (4) 1×10^{-4}M (5) 1×10^{-3}M (6) 5×10^{-3}M (7) 1×10^{-2} (8) 5×10^{-2}M

Assignments of α-proton resonances in sequence from the N-terminus are C, A, B, D.

those with multiple glycine units and the effective length being limited to 3 or 4 residues. An additional problem which limits the length of peptide which may be sequenced is a general broadening effect from unbound Gd^{3+} ions, the concentration of which increases with increasing total Gd^{3+} concentration, and with decreasing absolute concentration of ligand (fig. 4.1).

4D. SEQUENCE DETERMINATION BY ENHANCED SPIN LATTICE RELAXATION (T_1)

4D(i) Sequential relaxation for the Triglycine/Gd^{3+} system.

For a dipolar interaction, the paramagnetic contribution to T_1 and T_2 should be approximately equal. T_1 values for the α-protons would therefore be expected to show a sequential reduction depending on their distance from the bound Gd^{3+} ion.

The triglycine/Gd^{3+} system was used for initial experiments to test this prediction. It soon became apparent that a sequential reduction in T_1 values could be observed along the peptide chain at very low concentrations of Gd^{3+} (Table 4D I) which were insufficient to cause significant broadening of the α-proton resonances except that of the C-terminal protons (i.e. spectral resolution maintained during sequencing). The effect is so well defined that a single addition of Gd^{3+} is often all that is required to define a sequence.

TABLE 4D I

Reduction in T_1 for α-protons of Triglycine with Addition of Gd^{3+}

at pH 5.0 in 6M GuDCl measured at 270 MHz

Concentration of Gd^{3+} ($MX10^6$)	Concentration of peptides (M)	T_1 in secs.* of protons		
		N-terminal	Central	C-terminal
NIL	0.15	1.092	0.781	1.115
9.9	0.15	0.868 (21)	0.535 (31)	0.506 (55)
29.7	0.15	0.689 (37)	0.450 (42)	0.169 (85)
128.7	0.15	0.288 (74)	0.150 (81)	0.049 (96)

* Figures in brackets show the percentage reduction in T_1 caused by each addition of Gd^{3+}. The initial values are in substantial agreement with those obtained by Coates et al, (1973) whose values of 1.26, 0.91 and 1.24 secs. for N, Central and C-terminal α-protons respectively are slightly higher due to rigorous removal of oxygen prior to T_1 measurement. In the present work only comparative T_1 values before and after Gd^{3+} addition are needed and it was not necessary to remove dissolved oxygen.

4D(ii) Effective length of T_1 sequencing.

Polyglycines have been used as convenient models for assessment of the ultimate length of peptide which may be sequenced by this method. Table 4D II shows that the sequential relaxation effect can just be measured for all the protons in pentaglycine (with hexaglycine in 6M GuDCl it is not possible to resolve all resonances even at 270 MHz).

Undoubtedly the factor limiting the length of peptide which may be sequenced is spectrometer resolution. High field spectrometers (270 MHz) certainly help in extending the method to a pentapeptide but the National NMR Centre 270 MHz Brüker spectrometer used in this study has yet to produce the ultimate resolution with aqueous samples and it may well be that the limit extends beyond pentapeptides.

TABLE 4D II

Effect of Gd^{3+} on the T_1 values of ^1H nuclei in peptides in 6M GuDCl at pH 5.0

Peptide	Conc.^{n} (M)	Conc.^{n} Gd^{3+} ($M \times 10^6$)	T_1 secs. (Protons numbered from N-terminal)*				
			1	2	3	4	5
Tetraglycine	0.10	NIL	0.61	0.50	0.48	0.64	
		9.0	0.52 (15)	0.37 (26)	0.32 (34)	0.21 (67)	
Pentaglycine	0.025	NIL	0.52	0.46	0.43	0.47	0.42
		1.8	0.48 (9)	0.39 (14)	0.34 (20)	0.32 (32)	0.125 (70)

* Numbers in brackets refer to percentage reduction in T_1 after Gd^{3+} addition.

4D(iii) Extension to peptides other than polyglycines.

Spectrometer resolution is even more critical when glycine residues in the peptide chain are changed to other amino acids. However, it is still possible to sequence such peptides by T_1 measurement provided that at least one peak of the multiplet for each amino acid α-proton is resolved. Peptides so far successfully sequenced are shown in Table 4D III. The results on PheAspAlaSerVal highlight the problem of side chain carboxyl binding. In this case the aspartyl COO^- group is causing non sequential relaxation of the Asp α-proton and is undoubtedly causing some "premature" relaxation of both neighbouring α-protons. The pK of aspartyl side chain carboxyls is almost identical to the C-terminal group (Table 4DI) and hence it is not possible to overcome this problem by pH adjustment as was previously suggested (Crompton, 1973). It will be necessary to undertake some form of chemical modification or blocking of these groups if sequencing of peptides with aspartyl or glutamyl residues is to be successful.

TABLE 4D III

Peptide	Conc. (M)	Gd^{3+} ($Mx10^6$)	T_1 (secs). of protons numbered from N-terminus.*				
			1	2	3	4	5
IleIleIle	0.10	NIL	1.15	1.39	1.21		
		18	0.63 (45)	0.63 (55)	0.108 (91)		
PheAspAlaSerVal+	0.05	NIL	0.72**	0.80	0.72**	0.83	0.55
		5.5	0.35 (51)	0.121 (85)	0.35 (51)	0.105 (87)	0.043 (92)
GlyLeu	0.05	NIL	0.55	1.02	0.50	1.05	
		5.0	0.34 (38)	0.50 (51)	0.20 (60)	0.20 (81)	

* Numbers in brackets refer to percentage reduction in T_1 after addition of Gd^{3+}.

\+ T_1 values before and after Gd^{3+} addition for side chain protons are valine β CH, 0.56 and 0.167 (70).
Aspartic acid β - CH_2, 0.21 and 0.058 (97).

** These resonances completely overlap and hence the T_1 values are the averages of the two nuclei.

4E. PRACTICAL ASPECTS OF SEQUENCE DETERMINATION BY T_1 RELAXATION

4E(i) Amount of peptide required for NMR sequencing.

Any method of potential biological significance must of necessity by capable of analysing very small samples. All preceding NMR experiments have been undertaken using standard 5 mm tubes which require at least 20 mg. of peptide in order to maintain an absolute concentration at which optimum Gd^{3+} binding occurs. Attempts to use large volumes of very weak solutions only confirm the problem of poor lanthanide binding to the carboxyl group resulting in a large amount of general broadening. It was therefore deemed necessary to use some form of micro tube in which a small amount of peptide could be contained in a minimal volume of solvent. Two commercial designs of tube failed to give acceptable spectra, due it is thought to magnetic inhomogeneity problems caused by large amounts of solid glass or plastic in the magnetic field. An adjustable micro bulb tube (fig. 4.4) was designed to allow adjustable height of the sample bulb constructed from very thin glass. Acceptable NMR spectra (including T_1 measurements) were obtained from this tube when the bulb was immersed in D_2O contained in a standard NMR tube. The sample bulb is designed as a sliding fit inside the NMR tube to prevent wobble during spinning, and the capillary tube is a tight sliding fit through a standard NMR plastic tube cap which

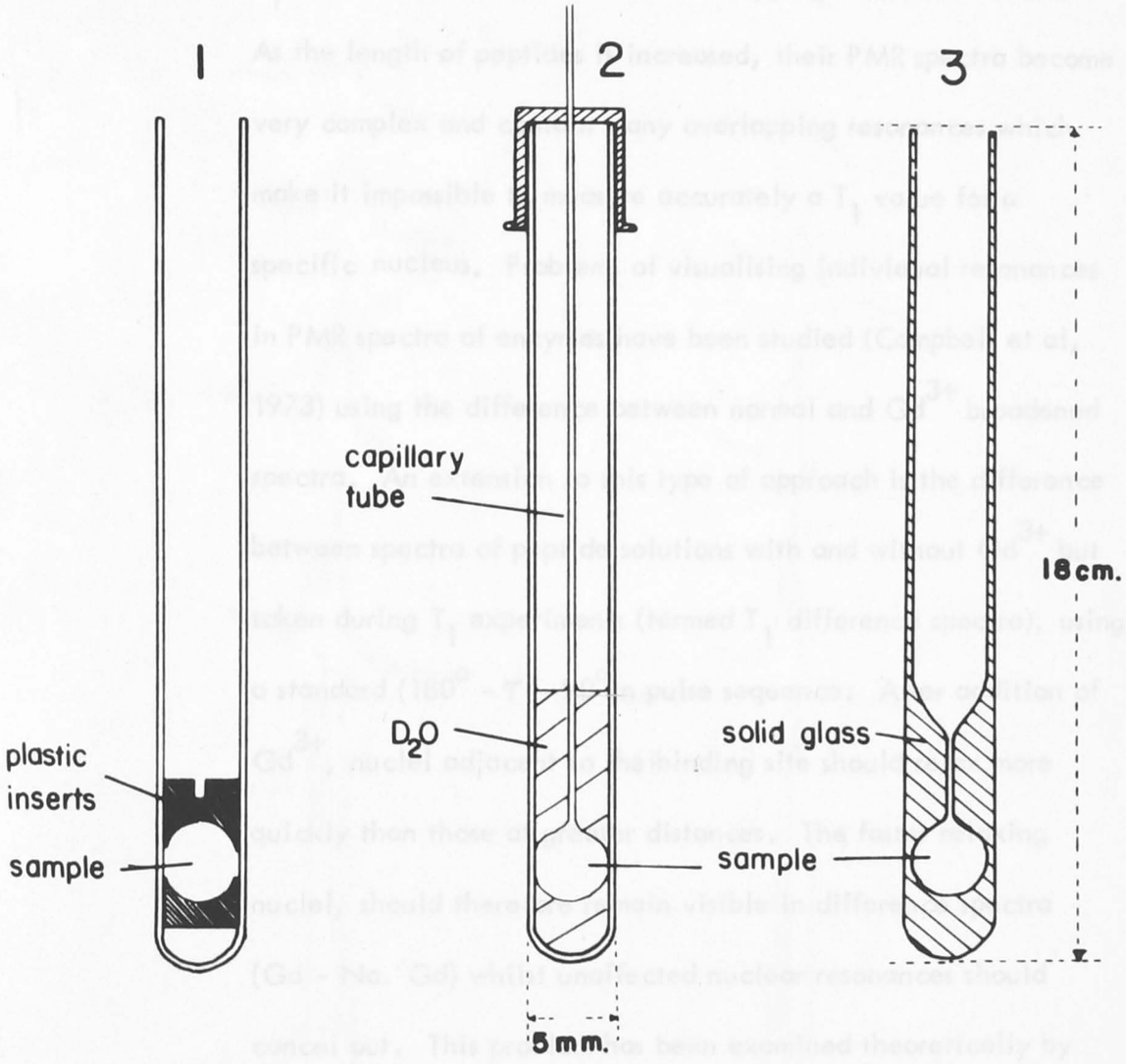


Fig. 4.4 Low volume NMR tubes. 1. Commercial design with plastic inserts. 2. Disposable micro bulb glass tube fabricated at ANU. 3. Commercial design in solid glass. The total volume of the bulb in type 2 is approx 20 microlitres.

allows height adjustment for the maximum signal. A comparison between the standard and micro tubes is shown in figure 4.5 for triglycine in a broadening experiment using Gd^{3+} at 100 MHz with a CW spectrometer.

4E(ii) T_1 measurement in cases where overlapping resonances occur.

As the length of peptides is increased, their PMR spectra become very complex and contain many overlapping resonances which make it impossible to measure accurately a T_1 value for a specific nucleus. Problems of visualising individual resonances in PMR spectra of enzymes have been studied (Campbell et al, 1973) using the difference between normal and Gd^{3+} broadened spectra. An extension to this type of approach is the difference between spectra of peptide solutions with and without Gd^{3+} but taken during T_1 experiments (termed T_1 difference spectra), using a standard $(180^\circ - \tau - 90^\circ)n$ pulse sequence. After addition of Gd^{3+}, nuclei adjacent to the binding site should relax more quickly than those at greater distances. The faster relaxing nuclei, should therefore remain visible in difference spectra (Gd - No. Gd) whilst unaffected nuclear resonances should cancel out. This problem has been examined theoretically by calculating curves for various relaxation times using a simple computer program (Appendix I). In fig. 4.6 the curves are shown for (a) the case where the peak height of the resonance in the presence of Gd^{3+} has decreased to $\frac{1}{3}$ of its original height and (b) where no decrease in height is noted. Plots of type (a)

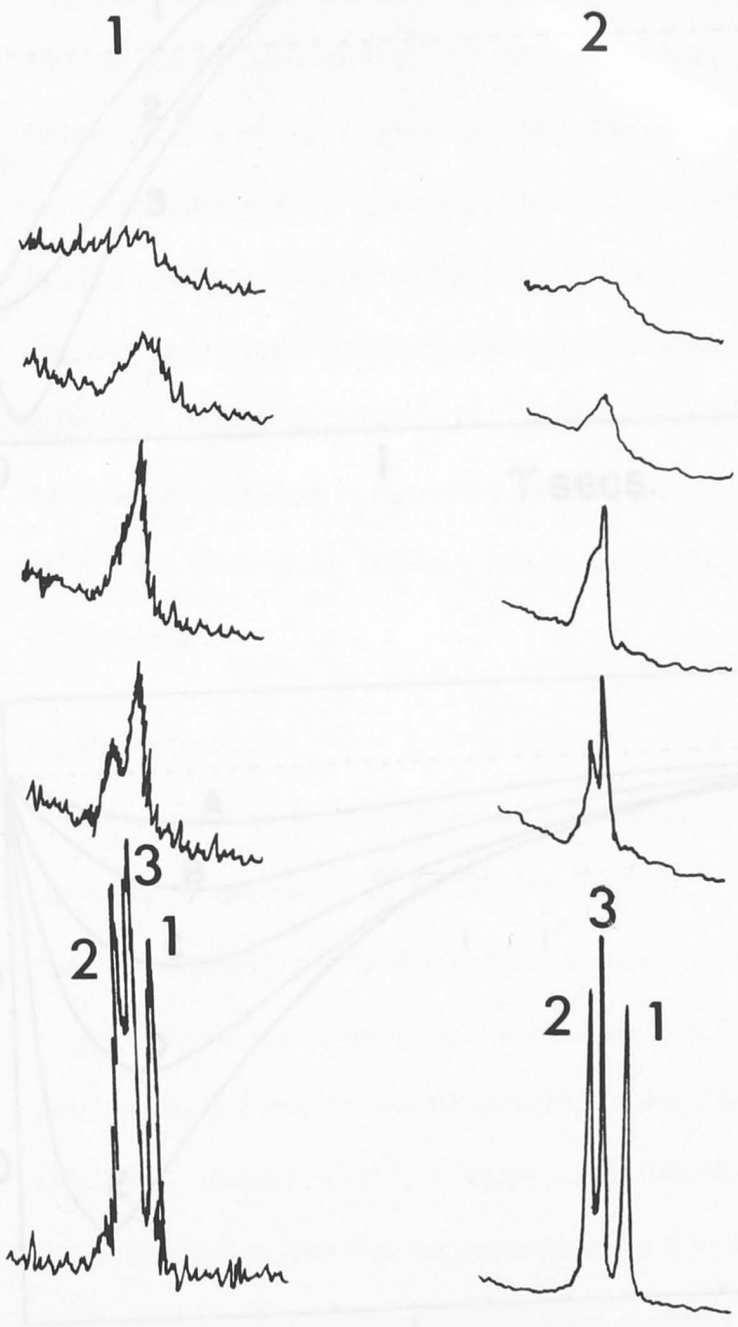


Fig. 4.5 Comparison of ^{1}H resonance broadening of triglycine with increasing concentration of Gd^{3+} in (1) a micro NMR tube containing 0.8 mg of triglycine and (2) a standard NMR tube containing 20 mg. Samples were dissolved in D_2O at pH 5.0 and spectra recorded on a 100 MHz continuous wave spectrometer.

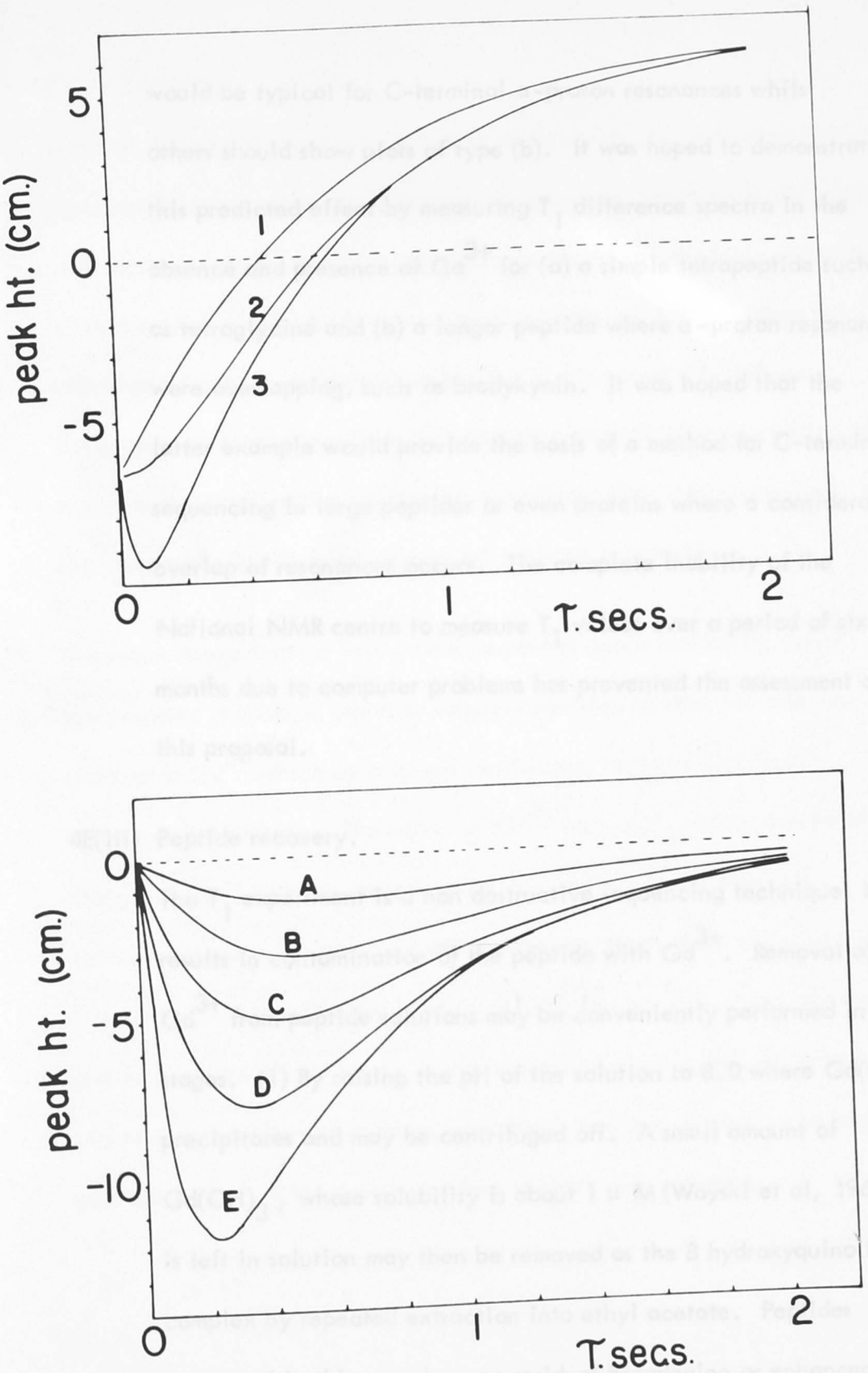


Fig. 4.6 Theoretical peak heights from T_1 difference spectra where initial height of peak for a 90° pulse is 10 cm at T_1 0.600 sec.

Top: Case where peak height is reduced to 50% of original height by the paramagnetic ion and T_1 values reduced to (1) 0.5, (2) 0.3 and (3) 0.1 secs.

Bottom: Peak heights not reduced but T_1 values reduced to (A) 0.5 (B) 0.4 (C) 0.3 (D) 0.2 (E) 0.1 secs.

would be typical for C-terminal α-proton resonances whilst others should show plots of type (b). It was hoped to demonstrate this predicted effect by measuring T_1 difference spectra in the absence and presence of Gd^{3+} for (a) a simple tetrapeptide such as tetraglycine and (b) a longer peptide where α-proton resonances were overlapping, such as bradykynin. It was hoped that the latter example would provide the basis of a method for C-terminal sequencing in large peptides or even proteins where a considerable overlap of resonances occurs. The complete inability of the National NMR centre to measure T_1 values over a period of six months due to computer problems has prevented the assessment of this proposal.

4E(iii) Peptide recovery.

The T_1 experiment is a non destructive sequencing technique, but results in contamination of the peptide with Gd^{3+}. Removal of Gd^{3+} from peptide solutions may be conveniently performed in two stages. (1) By raising the pH of the solution to 8.0 where $Gd(OH)_3$ precipitates and may be centrifuged off. A small amount of $Gd(OH)_3$, whose solubility is about 1 μ M (Woyski et al, 1963) is left in solution may then be removed as the 8 hydroxyquinoline complex by repeated extraction into ethyl acetate. Peptides recovered in this way show no residual broadening or enhanced relaxation of the C-terminal α-proton resonances.

4F. CONCLUSIONS

It is not thought likely that broadening experiments (T_2 measurements) will be very useful for sequence determination except in certain special cases where resonances from each α-proton in the peptide are widely separated. Even then it is likely to be limited to tripeptides, which can often be sequenced solely by resonance pH shifts. If a successful broadband decoupling apparatus could be built to allow simultaneous decoupling of all α - β protons during the broadening experiment then it is conceivable that the length of peptide to be sequenced could be increased.

Sequence determination by T_1 measurement is again limited by overall spectrometer resolution but only requires addition of very low concentrations of Gd^{3+} which usually are insufficient to cause significant resonance broadening. The effect of Gd^{3+} ions on T_1 values can be sensed along the chain of a pentapeptide but cannot be used for peptides containing aspartyl or glutamyl residues except when these occur at the C-terminus, when additonal binding of Gd^{3+} on the side chain carboxyl groups could be advantageous. The possibility of using this technique for sequencing a number of residues from the C-terminus of medium to long peptides by a T_1 difference procedure appears very promising.

CHAPTER 5

N-TERMINAL SEQUENCING IN PEPTIDES

5A. INTRODUCTION

The biological significance of copper complexes of peptides and proteins has resulted in an intense interest in their structure and properties (Peisach et al, 1966). A major problem is in the location of the copper atom in the protein (Gould et al, 1973). Compared to iron heme proteins, where at lease four of the six ligands are easily identified after isolation and characterisation of the prosthetic group, copper atoms in such complexes may have three to six ligands all donated by peptide chain residues (Gould et al, 1973) and hence identification of these ligands eventually requires a direct study on the protein. Before attempting to study such a complicated system much background data can be obtained from an examination of the structures and properties of model copper peptide complexes. It is in this field that NMR has proved particularly useful. The observation that paramagnetic Cu^{2+} ions cause selective perturbations in the NMR spectra of tripeptides (Li et al, 1962) has formed the basis of the following investigation to assess the usefulness of copper peptide complexes for sequence determination.

It is important to gain an understanding of the effect of Cu^{2+} ions on the NMR solution spectra of peptides and to aquire much basic data about the system so that optimum conditions for sequence determination may be selected. In particular it is essential to understand how copper binds to peptides, which groups co-ordinate and to what extent solution parameters (peptide or metal concentration, pH, etc.) affect the type and distribution of complexes. Of equal importance is the necessity to ascertain the type of mechanism responsible for the magnetic perturbations, i.e. is the interaction between paramagnetic ions and peptide nuclei predominantly dipolar, scalar, or governed by some exchange phenomenon.

5B. COPPER COMPLEXES OF PEPTIDES

5B(i) Co-ordination sites.

Some of the earliest X-ray crystallographic studies on copper triglycine complexes (Freeman et al, 1964) showed that in crystals obtained from equimolar neutral solutions, each copper atom was bound to two different peptide molecules and each peptide to two different copper atoms. This results in the formation of infinite chains of composition $(Cu_2A_2^{2+})n$, but from alkaline solution crystals of the dimeric species $Cu_2H_{-4}A_2^{2-}$ were obtained (Freeman et al, 1965). In solution the polymeric species of the crystalline state break down to smaller molecules of the type Cu_2A^{3+}, CuA^{+}, CuA_2, and the distribution of these species

under varying solution conditions becomes very complex. In such cases where competing equilibria are present there are many factors which determine the extent to which a particular complex predominates, such as the relative and absolute concentration of metal ions and ligand, pK values of ligand, pH of solution, and the stability constants of each metal complex.

Solution studies (Dobbie et al, 1955) indicated the presence of a series of mononuclear complexes of copper and triglycine with co-ordination at the N-terminal group with additional and progressive co-ordination to each deprotonated peptide bond as the pH was raised. These complexes can be represented by the series $CuHqA^{(1+q)+}$ where q=1, 0, -1, -2. In addition the series $CuHqA_2^{q+}$ with q=1, or -2 was also thought to exist (Dobbie et al, 1955). These results have since been partially confirmed by Koltun et al, (1963), and Kim and Martell (1966). More recently Osterberg and Sjoberg (1968) and Perrin et al, (1972) carried out extensive studies to measure stability constants of these species. In particular Osterberg and Sjoberg (1968) confirmed the presence of three major groups of species in the system Cu^{2+} /triglycine. At low metal and low triglycine concentration the mononuclear species $CuHqA^{(1+q)+}$ predominates (q=1, 0, -1, -2). At concentrations above 10mM, in the region where most NMR work is performed, there is strong evidence for the series $CuHqA_2^{q+}$ with q=0, -1, -2, and $Cu_2HqA^{(2+q)+}$ with q=1, 0, -2, -4. Very slight evidence existed for the presence of traces of $CuH_2A_2^{2+}$ and $Cu_2H_2A_2^{4+}$.

5B(ii) Structures of complexes.

In NMR experiments where the concentration of ligand is always in excess over that of copper, it has been shown that the major solution species for copper triglycine complexes consist almost entirely of the two groups $CuHqA_2^{q+}$ and $CuHqA^{(1+q)+}$ (Osterberg and Sjoberg, 1968). Some structures have already been proposed for these complexes and these are shown in fig. 5.1 (I) (II) (III) together with logical additions to each series. A further possible structure to the mononuclear series not shown in this figure is CuH_2A^{3+} where a single copper atom is bound directly to the ionised C-terminal carboxyl group, leaving an unbound protonated N-terminal amino group. It is important to note at this stage that these structures have been obtained from potentiometric studies at higher copper to ligand ratios than those used in NMR work, and it is conceivable that in cases where the ligand concentration exceeds that of copper by many orders of magnitude, other species might exist. Complexes of the type CuA_3 or CuA_4 recently found in the glycine copper system (Beattie et al, 1976) where three or four ligands are bound only by their N-terminal amino group to a single copper atom, seem most likely.

Mononuclear species $CuH_qA^{(1+q)+}$

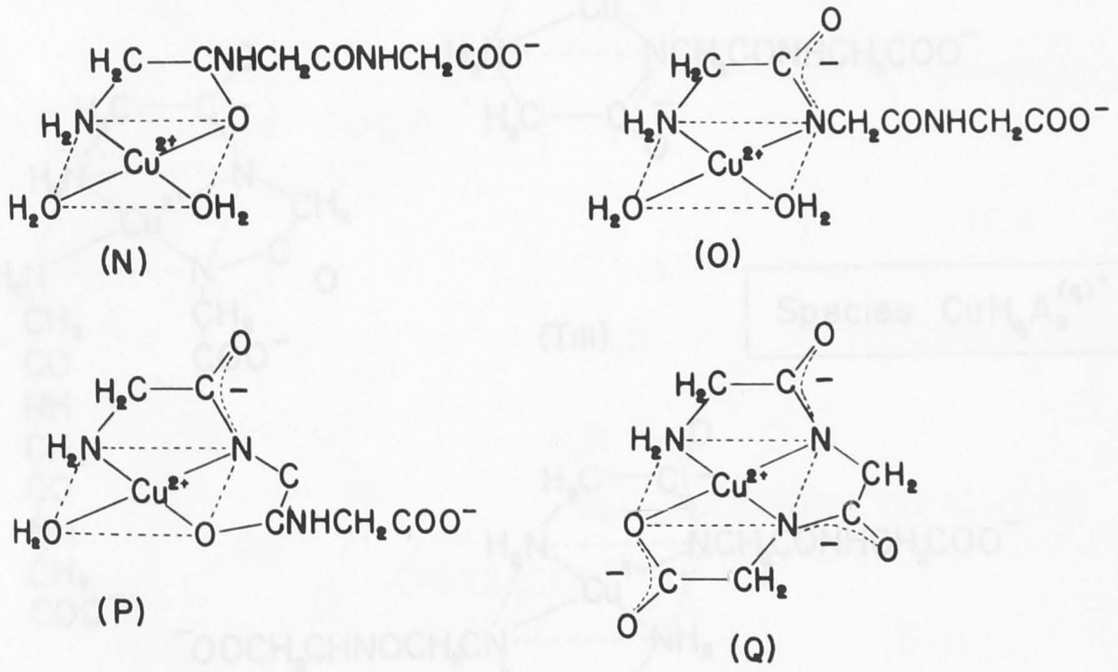


Fig. 5.1 (I) Structures of copper triglycine complexes for the series $CuHqA^{(1+q)+}$

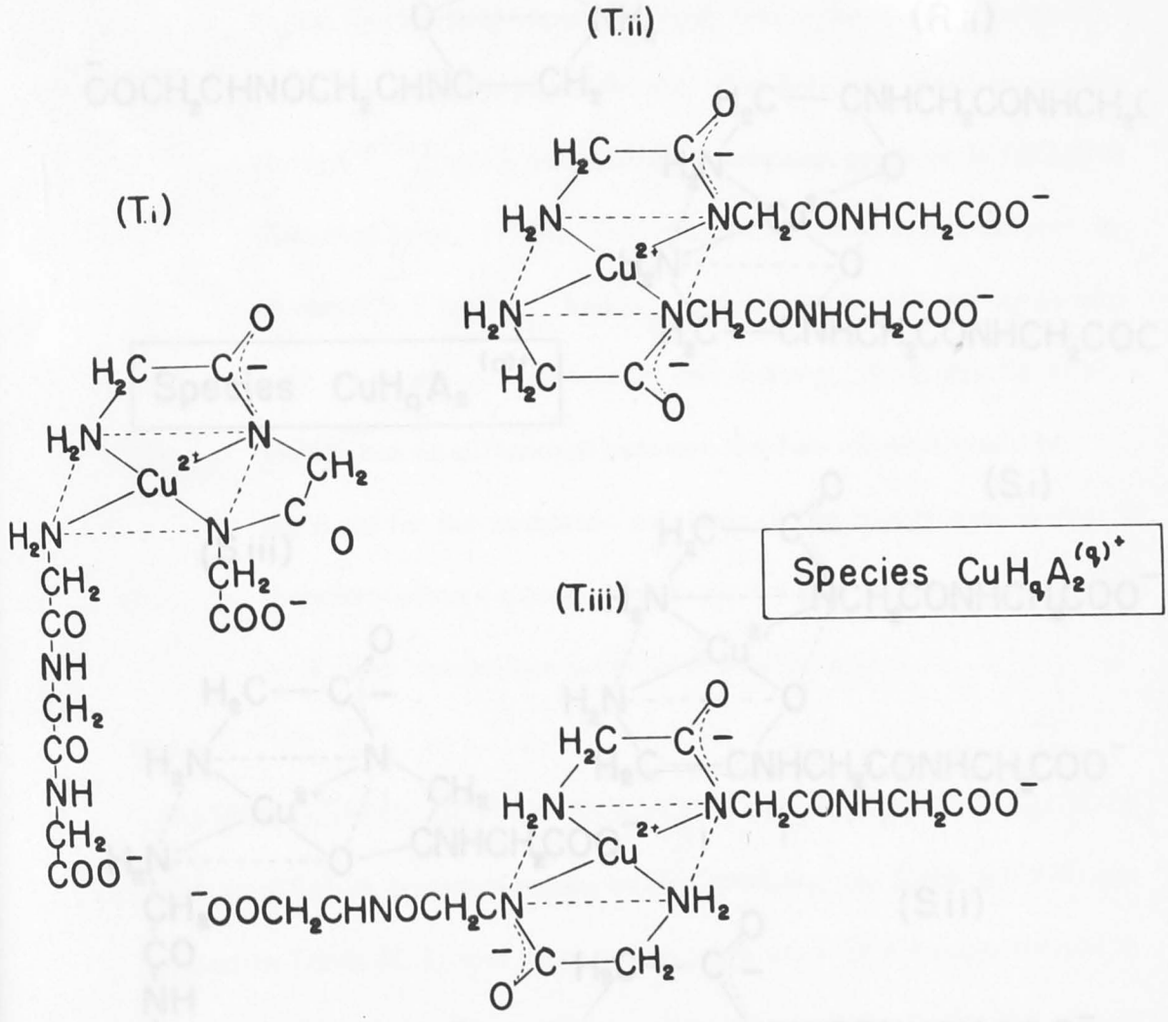


Fig. 5.1 (II) Structures for species $CuHqA2^{q+}$

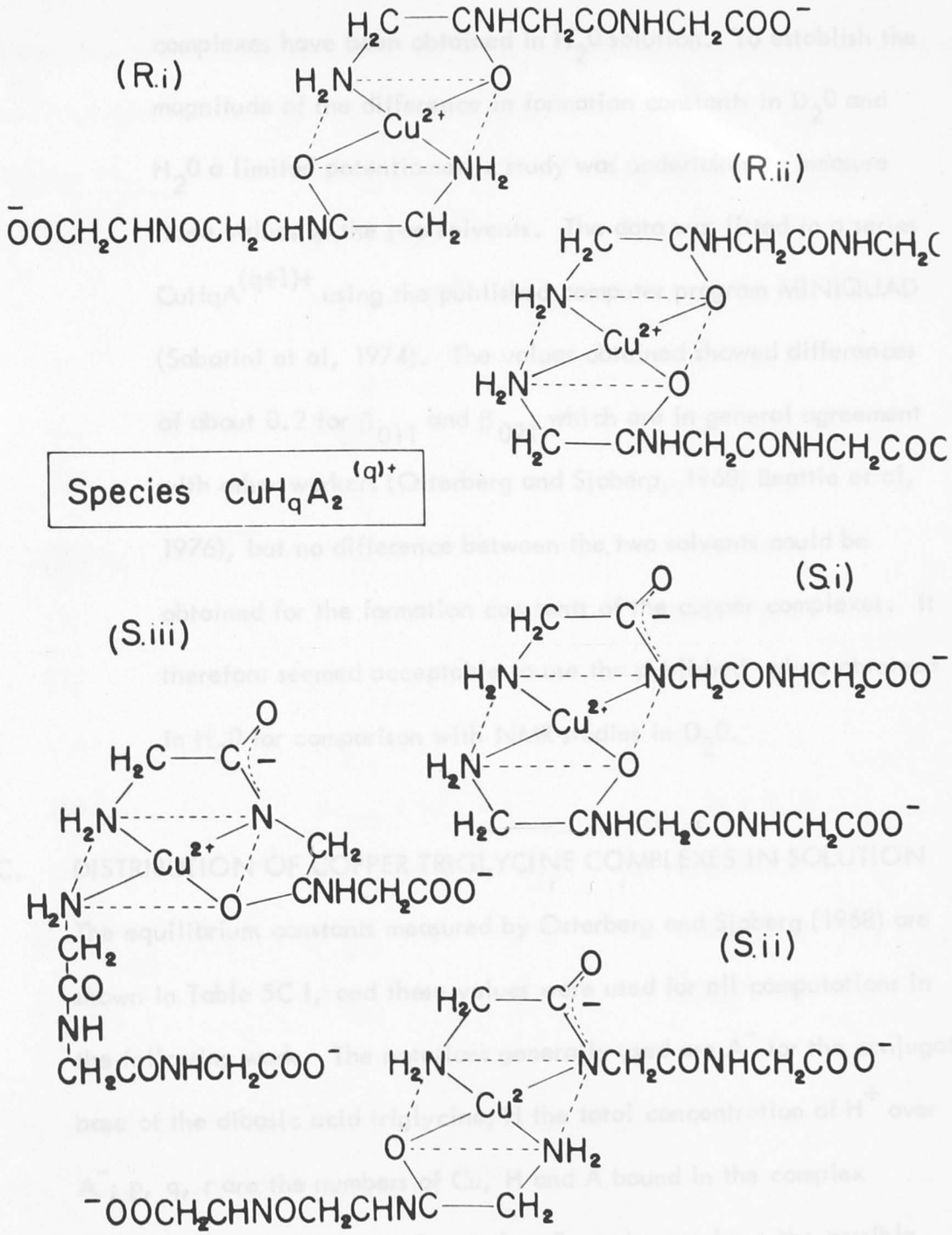


Fig. 5.1 (III) Structures for species $CuHqA_2^{q+}$

5B(iii) Equilibrium constants in D_2O

Most published values of formation constants for copper triglycine complexes have been obtained in H_2O solution. To establish the magnitude of the difference in formation constants in D_2O and H_2O a limited potentiometric study was undertaken to measure these values in the two solvents. The data was fitted to a series $CuHqA^{(q+1)+}$ using the published computer program MINIQUAD (Sabatini et al, 1974). The values obtained showed differences of about 0.2 for β_{011} and β_{021} which are in general agreement with other workers (Osterberg and Sjoberg, 1968; Beattie et al, 1976), but no difference between the two solvents could be obtained for the formation constants of the copper complexes. It therefore seemed acceptable to use the published values obtained in H_2O for comparison with NMR studies in D_2O.

5C. DISTRIBUTION OF COPPER TRIGLYCINE COMPLEXES IN SOLUTION

The equilibrium constants measured by Osterberg and Sjoberg (1968) are shown in Table 5C I, and these values were used for all computations in the following work. The notations generally used are A^- for the conjugate base of the dibasic acid triglycine; H the total concentration of H^+ over A^-; p, q, r are the numbers of Cu, H and A bound in the complex $Cu_pH_qA_r$ where p is always 1, r is 1 or 2, and q can have the possible values of 1, 0, -1, -2. The negative values indicate displacement of protons from the ligand which do not normally dissociate in the absence of complexing metal ions.

TABLE 5C I

Formation constants for triglycine/Cu^{2+} complexes from Osterberg et al, (1968). These values were obtained over a metal ion concentration range of 1 to 100 mM and triglycine concentration of 2 to 250 mM.

Species No.*	pqr	Log equilibrium constant
(N)	$\beta_{111} \times 10^{-10}$	1.34 ± 0.05
(O)	$\beta_{101} \times 10^{-5}$	4.5 ± 0.2
(P)	$\beta_{1-11} \times 10$	8.0 ± 0.4
(Q)	$\beta_{1-21} \times 10^{7}$	1.5 ± 0.1
(X)	$\beta_{122} \times 10^{-19}$	0.9 ± 0.3
(R)	$\beta_{102} \times 10^{-10}$	1.39 ± 0.07
(S)	$\beta_{1-12} \times 10^{-3}$	8.0 ± 0.5
(T)	$\beta_{1-22} \times 10^{5}$	1.50 ± 0.15
(U)	$\beta_{212} \times 10^{-17}$	1.9 ± 0.3
(V)	$\beta_{202} \times 10^{-13}$	1.24 ± 0.08
(W)	$\beta_{2-22} \times 10^{-1}$	3.0 ± 0.3
pK_1	β_{011}	8.547
pK_{1+2}	β_{021}	12.259

The overall equilibrium constants β_{pqr} are for the general reaction

$$pCu^{2+} + qH^{+} + rA^{-} \rightleftharpoons CupHqAr^{(2p+q-r)+}$$

* Letters refer to the species noted in the distribution curves (fig. 5.2).

5C(i) Effect of pH on distributions.

To understand how pH affects the distribution of triglycine copper complexes a computer program SIAS (Sylva, 1975, Appendix I) was used to compute the distributions at varying pH and at concentrations of triglycine and copper used in NMR studies. The results (fig. 5.2) emphasise the following points:

(a) Free Cu^{2+} can exist up to about pH 5.

(b) The predominant species particularly at higher pH is clearly $CuHqA_2^{q+}$ with q varying from 0, -1, -2 as the pH is raised.

(c) Lesser amounts of $CuHqA^{(1+q)+}$ are present at lower pH with values with q varying from 1, 0, -1, -2 as the pH is raised.

(d) Where Cu^{2+} is in large excess over triglycine, then the series $Cu_2HqA^{(2+q)+}$ is not important.

5C(ii) Effect of ligand concentration.

Computed distributions at a fixed Cu^{2+} concentration of 1×10^{-5}M and a triglycine concentration from 0.001 to 0.5M were obtained (fig. 5.2) and clearly demonstrate the controlling effect of absolute ligand concentration. As the ligand concentration is reduced the range of pH at which multiple species occur is considerably reduced until at a ligand concentration of 0.001M the distribution is almost solely between free Cu^{2+} and $CuH_{-2}A_2^{2+}$.

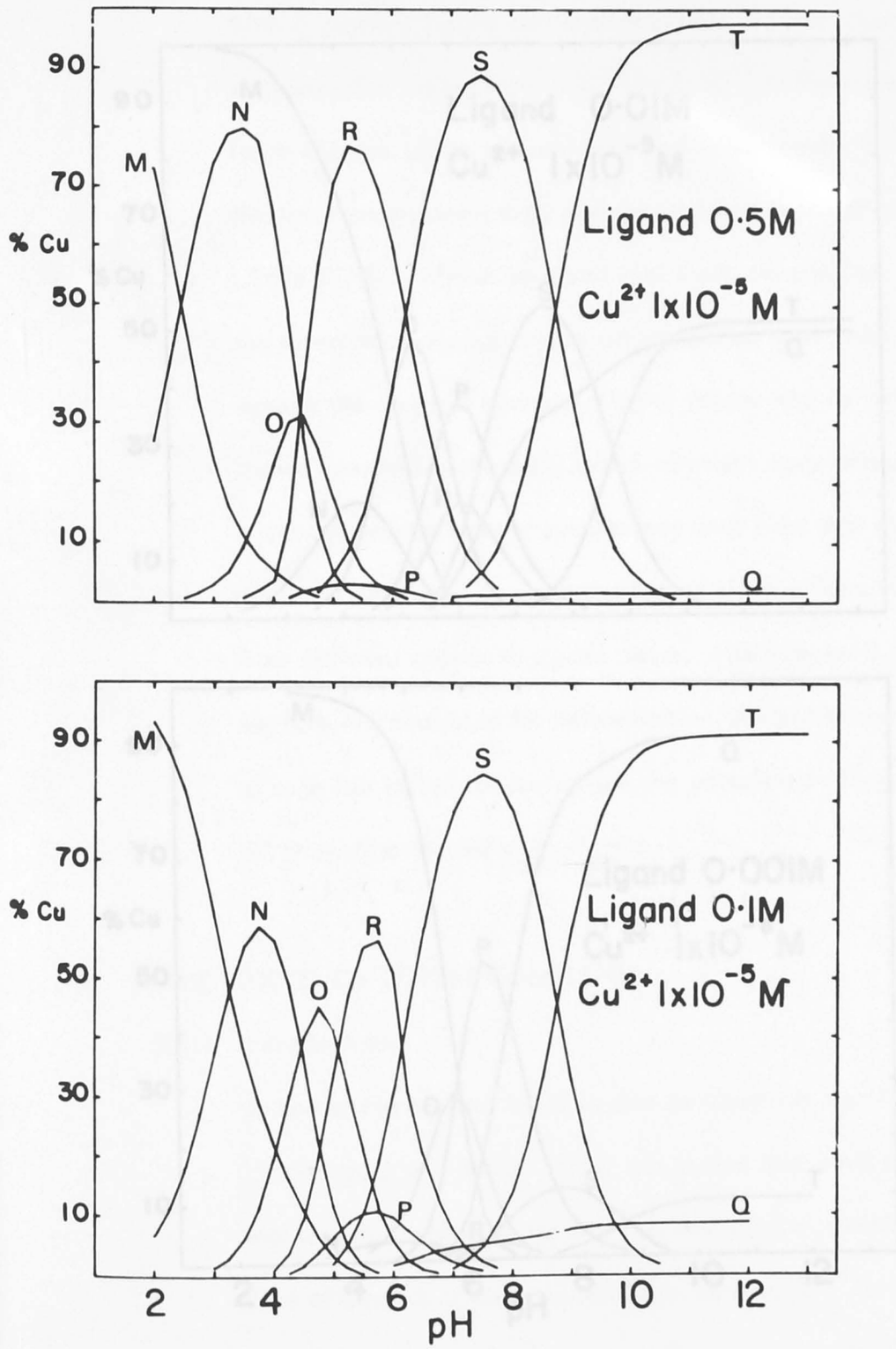


Fig. 5.2 pH distributions of the copper triglycine species proposed by Osterberg et al[10]. Graphs show the percentage of total copper in the species.
(M) Cu^{2+} (N) 1,1,1, (O) 1,0,1, (P) 1,-1,1, (Q) 1,-2,1, (R) 1,0,2, (S) 1,-1,2, (T) 1,-2,2, where each number refers to pqr in the complex CupHqAr

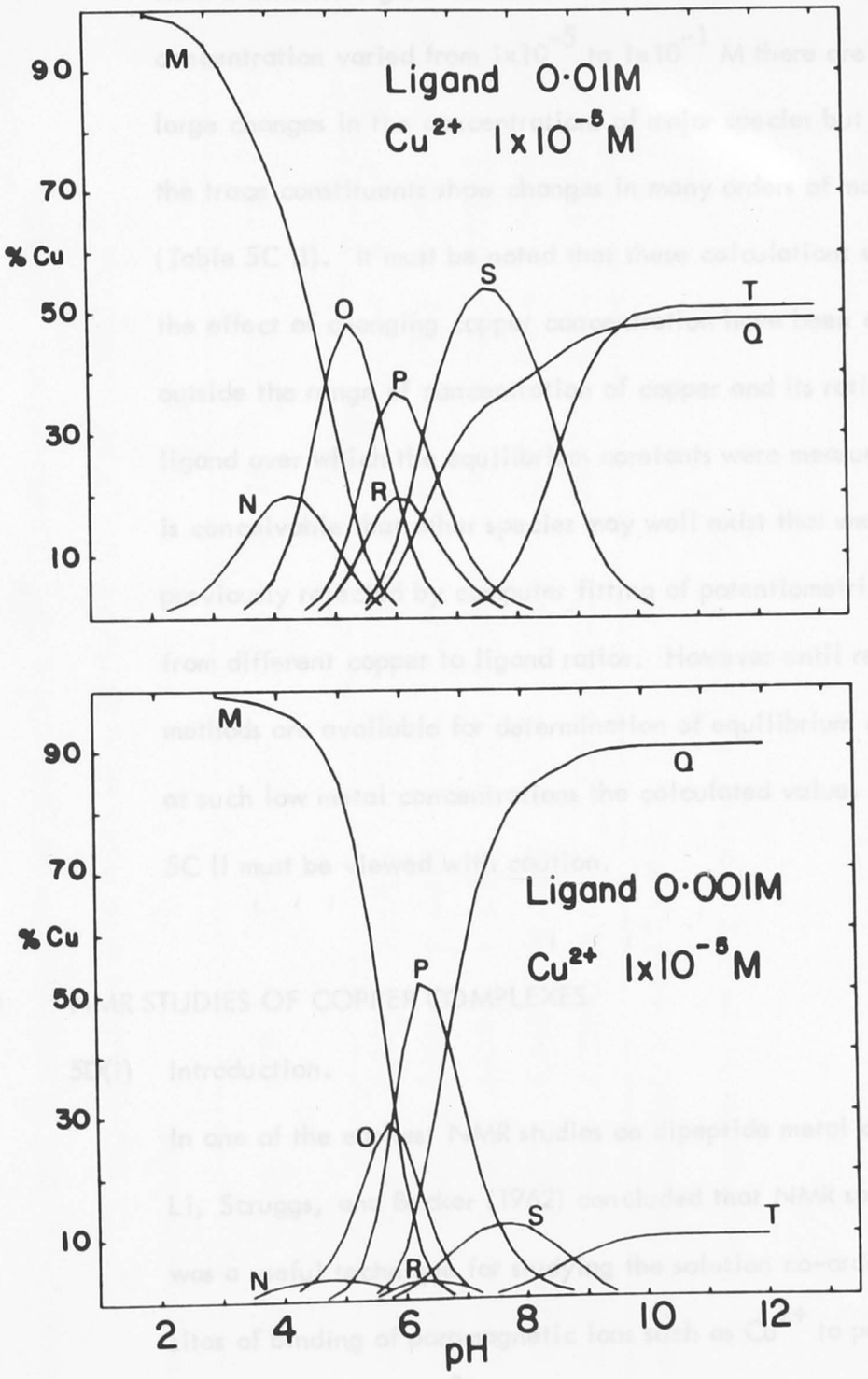


Fig. 5.2 (continued)

5C(iii) Effect of copper concentration.

With a constant ligand concentration of 0.2M and a total copper concentration varied from 1×10^{-5} to 1×10^{-1} M there are no large changes in the concentrations of major species but some of the trace constituents show changes in many orders of magnitude (Table 5C II). It must be noted that these calculations showing the effect of changing copper concentration have been computed outside the range of concentration of copper and its ratio to ligand over which the equilibrium constants were measured. It is conceivable that other species may well exist that were previously rejected by computer fitting of potentiometric data from different copper to ligand ratios. However until reliable methods are available for determination of equilibrium constants at such low metal concentrations the calculated values in Table 5C II must be viewed with caution.

5D NMR STUDIES OF COPPER COMPLEXES

5D(i) Introduction.

In one of the earliest NMR studies on dipeptide metal complexes Li, Scruggs, and Becker (1962) concluded that NMR spectroscopy was a useful technique for studying the solution co-ordination sites of binding of paramagnetic ions such as Cu^{2+} to peptides. They noted that at Cu^{2+} concentrations as low as 1×10^{-5} M the N-terminal α-proton resonances of diglycine, glycylalanine, and glycylproline broadened and disappeared completely from their PMR spectra. No effect was observed on the C-terminal

TABLE 5C II

VARIATION IN SPECIES DISTRIBUTION OF TRIGLYCINE/Cu^{2+} COMPLEXES WITH VARYING Cu^{2+} CONCENTRATION WITH FIXED pH AND TRIGLYCINE CONCENTRATION

(A) pH 5.0 Ligand concentration 0.2M

Cu^{2+} conc.^{n}(M)	(M) Free Cu^{2+}	(N)	(O)	(P)	(Q)	(X)	(R)	(S)	(T)	(U)	(V)	(W)
1×10^{-5}	1.21E+0	8.80E+0	3.07E+01	5.04E+0	9.45E+2	1.70E-3	1.11E-2	2.50E-4	3.44E-9	5.11E+1	3.10E+0	5.48E-4
1×10^{-4}	1.21E+0	8.80E+0	3.06E+1	5.04E+0	9.44E-2	1.70E-2	1.11E-1	2.50E-3	3.43E-8	5.10E+1	3.10E+0	5.47E-4
1×10^{-3}	1.21E+0	8.72E+0	3.04E+1	5.00E+0	9.37E-2	1.67E-1	1.10E+0	2.46E-2	3.38E-7	5.03E+1	3.05E+0	5.38E-4
1×10^{-2}	1.20E+0	8.13E+0	2.83E+1	4.66E+0	8.73E-2	1.45E+0	9.51E+0	2.13E-1	2.93E-6	4.38E+1	2.66E+0	4.69E-4
1×10^{-1}	1.79E+0	5.67E+0	1.98E+1	3.25E+0	6.09E-2	7.05E+0	4.63E+1	1.04E+0	1.43E-5	1.43E+1	8.66E-1	1.53E-4

Column heading refer to species listed in Table 5C I and fig. 5.1.

Values show the % of copper in each species (in scientific notation, i.e. 1.5E+2 = 1.5×10^{2})

TABLE 5C II

(continued)

(B) pH 7.0

Cu^{2+} conc.^{n}(M)	(M) Free Cu^{2+}	(N)	(O)	(P)	(Q)	(X)	(R)	(S)	(T)	(U)	(V)	(W)
1×10^{-5}	3.03E-5	2.26E-4	7.87E-2	1.29E+0	2.43E+0	1.12E-10	7.34E-8	1.65E-5	2.26E-6	1.34E+1	8.14E+1	2.44E+0
1×10^{-4}	3.04E-5	2.26E-4	7.88E-2	1.30E+0	2.43E+0	1.12E-9	7.35E-7	1.65E-4	2.27E-5	1.34E+1	8.14E+1	1.44E+0
1×10^{-3}	3.09E-5	2.28E-4	7.94E-2	1.31E+0	2.45E+0	1.14E-8	7.48E-6	1.68E-3	2.31E-4	1.34E+1	8.13E+1	1.44E+0
1×10^{-2}	3.71E-5	2.49E-4	8.68E-2	1.43E+0	2.68E+0	1.36E-7	8.94E-5	2.01E-2	2.76E-3	1.34E+1	8.10E+1	1.43E+0
1×10^{-1}	1.25E-3	1.26E-3	4.40E-1	7.23E+0	1.36E+1	3.50E-5	2.29E-2	5.15E+0	7.07E-1	1.02E+1	6.16E+1	1.09E+0

TABLE 5C II

(continued)

(C) pH 9.0

Cu^{2+} conc.^{n}(M)	(M) Free Cu^{2+}	(N)	(O)	(P)	(Q)	(X)	(R)	(S)	(T)	(U)	(V)	(W)
1×10^{-5}	1.82E-10	1.25E-11	1.37E-6	7.15E-3	4.24E+0	1.08E-22	2.24E-17	5.03E-10	6.91E-6	7.58E-3	1.45E+1	8.12E+1
1×10^{-4}	1.83E-10	3.62E-10	1.26E-5	2.08E-2	3.89E+0	2.88E-19	1.89E-14	4.24E-8	5.82E-5	5.72E-2	3.47E+1	6.13E+1
1×10^{-3}	1.86E-10	3.65E-10	1.27E-5	2.09E-2	3.92E+0	2.93E-18	1.92E-13	4.31E-7	5.92E-4	5.72E-2	3.47E+1	6.13E+1
1×10^{-2}	2.23E-10	3.99E-10	1.39E-5	2.29E-2	4.29E+0	3.50E-17	2.30E-12	5.15E-6	7.08E-3	5.70E-2	3.46E+1	6.11E+1
1×10^{-1}	8.56E-9	2.18E-9	7.60E-5	1.25E-1	2.34E+1	1.04E-14	6.85E-10	1.54E-3	2.11E+0	4.43E-2	2.69E+1	4.74E+1

TABLE 5C II

(continued)

(D) pH 11.0

Cu^{2+} conc.^{n}(M)	(M) Free Cu^{2+}	(N)	(O)	(P)	(Q)	(X)	(R)	(S)	(T)	(U)	(V)	(W)
1×10^{-5}	1.56E-14	4.17E-16	1.45E-0	2.39E-4	4.48E+0	3.81E-30	2.50E-23	5.62E-13	7.72E-7	8.86E-6	5.38E-1	9.50E+1
1×10^{-4}	1.56E-14	4.21E-16	1.47E-9	2.39E-4	4.48E+0	3.82E-29	2.51E-22	5.63E-12	7.73E-5	8.86E-6	5.38E-1	9.50E+1
1×10^{-3}	1.59E-14	4.21E-16	1.47E-9	2.41E-4	4.52E+0	3.89E-28	2.55E-21	4.72E-11	7.86E-4	8.86E-6	5.38E-1	9.49E+1
1×10^{-2}	1.91E-14	4.60E-16	1.60E-9	2.63E-4	4.94+0	4.64E-27	3.04E-20	6.83E-10	9.38E-3	8.82E-6	5.35E-1	9.45E+1
1×10^{-1}	6.37E-13	2.32E-15	8.09E-9	1.33E-3	2.49+1	1.18E-24	7.76E-18	1.74E-7	2.39E+0	6.74E-6	4.09E-1	7.23E+1

α-protons. They inferred that the sites of binding to the Cu^{2+} ion were the terminal amino group and the adjacent peptide amide group because the α-proton resonance affected by the paramagnetic ion was located between these two sites. The work was not extended to higher copper concentrations. Later studies by Li and Tang, (1964) revealed a similar effect for some tripeptides and in particular for triglycine. It was not until the work of Kim and Martell, (1969) that the effect of higher concentrations of copper and nickel on the PMR spectra of tri and tetrapeptides was studied. They concluded that the order in which α-proton resonances were broadened by Cu^{2+} suggested the manner in which metal ions co-ordinate under changing solution conditions. They found that at high pH the α-proton resonances broaden selectively from N- to C-terminus for diglycine and tetraglycine but not in the case of triglycine where both central and C-terminal α-proton signals broadened simultaneously. No attempt was made to relate these observations to the distribution of solution species except in the case of triglycine. For this peptide the non sequential resonance broadening was explained by strong N-terminal co-ordination and simultaneous but weaker carboxylate binding. Very slow amide deprotonation allows only a very weak copper to peptide bond at this site. Assuming this to be a reasonable explanation for triglycine it then seems improbable that amide deprotonation with subsequent copper binding is a logical explanation for the broadening sequence of tetraglycine

resonances in which the C-terminal α-proton resonance is the last to broaden. In view of the many factors which affect the distribution of species in solution, which have been discussed earlier in this chapter, the work of Kim and Martell, (1969) can be severly criticised on the following experimental detail:

(a) Control of ligand concentration.

About 5 to 10% w/w solutions were used. No accurate concentrations were prepared nor was there any attempt to maintain a constant peptide concentration between or during experiments.

(b) Control of pH.

In the preparation of acidic or alkaline solutions slightly more than equivalent amounts of NaOD or DCl were added. The exact pH or pD of each experiment was not recorded. Because Cu^{2+} ions are displacing protons from peptide bonds each addition of copper causes the pH to drop and a corresponding pH adjustment is necessary to maintain a constant value.

(c) Control of copper concentration.

"Small amounts" of the metal chloride were added to each solution, and if precipitation occured, only the supernatant liquid was used. Hence there was no knowledge of the exact copper concentration or of its ratio to peptide.

5D(ii) Experiments to study the causes of resonance line broadening by Cu^{2+}.

The copper triglycine system was chosen as a model system to gain an understanding of the types of complexes and relaxation mechanisms responsible for the observed line broadening in the NMR spectra. Such basic information is necessary to aid the selection of suitable conditions for the sequence determination of di, tri, and longer peptides.

The PMR spectra of triglycine consists of three singlet resonances from the three α-protons. The resonances of the C- and N-terminal α-protons undergo a downfield shift during protonation of the COO^- and $-NH_2$ groups respectively whilst the central α-proton resonance shows only a very slight downfield shift during each protonation of adjacent terminal groups.

A. Effect of copper concentration at varying pH.

Curves of resonance peak height (normalised) plotted against copper concentration at defined pH were obtained from experiments where the pH and ligand concentration were maintained at constant values during the addition of copper. Selected graphs in which curves are numbered sequentially from N- to C-terminus with Cu^{2+} addition are shown in fig. 5.3. Over the range of copper concentration normally required to broaden all the ^1H resonances of triglycine no detectable change in chemical

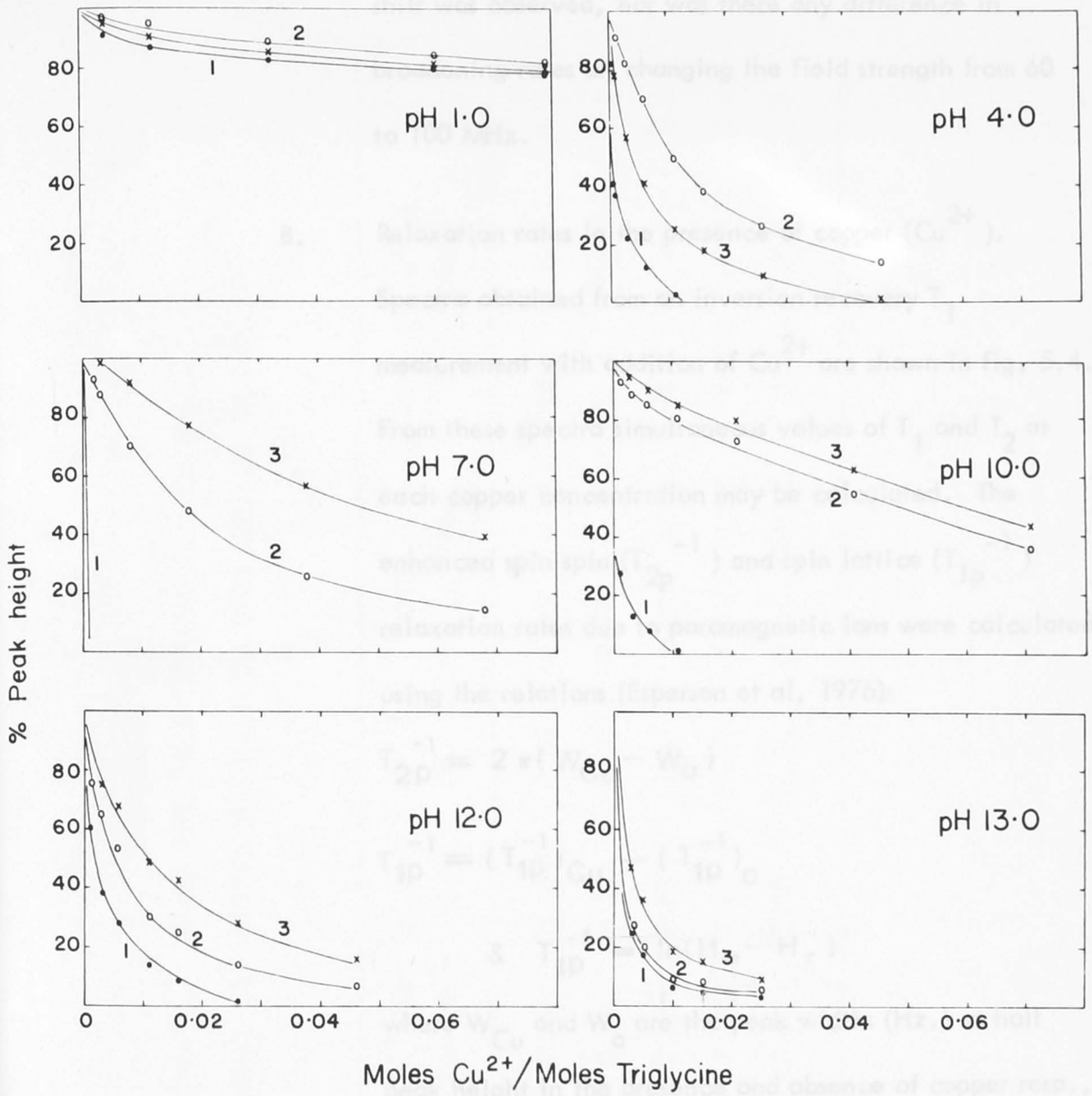


Fig. 5.3 Effect of the concentration of Cu^{2+} on the α-proton resonance peak heights of triglycine at varying pH. Numbers on each graph refer to the α-protons of triglycine numbered sequentially from the N-terminus.

shift was observed, nor was there any difference in broadening rates on changing the field strength from 60 to 100 MHz.

B. Relaxation rates in the presence of copper (Cu^{2+}).

Spectra obtained from an inversion recovery T_1 measurement with addition of Cu^{2+} are shown in fig. 5.4. From these spectra simultaneous values of T_1 and T_2 at each copper concentration may be calculated. The enhanced spin spin (T_{2p}^{-1}) and spin lattice (T_{1p}^{-1}) relaxation rates due to paramagnetic ions were calculated using the relations (Espersen et al, 1976):

$$T_{2p}^{-1} = 2\pi(W_{Cu} - W_o)$$

$$T_{1p}^{-1} = (T_{1p}^{-1})_{Cu} - (T_{1p}^{-1})_o$$

$$\& \quad T_{1p}^{-1} = \ln(H_\infty - H_\tau)$$

where W_{Cu} and W_o are the peak widths (Hz.) at half peak height in the presence and absence of copper resp., H_∞ and H_τ are the peak heights with delay times of ∞ and τ respectively and $(T_{1p}^{-1})_{Cu}$ and $(T_{1p}^{-1})_o$ are the inverse spin lattice relaxation times in the presence and absence of copper. Values obtained for (T_{1p}) Cu and $(T_{1p})_o$ from a computer least squares analysis (Appendix I) of the plot $^1/T_{1p}$ vs. $\ln(H_\alpha - H_\tau)$ and the ratios of T_{1p}/T_{2p} for protons of triglycine are shown in Table 5D I.

TABLE 5D 1

T_{1P}^{-1} AND T_{2P}^{-1} VALUES FOR THE N-TERMINAL (N), CENTRAL, AND C-TERMINAL (C) PROTONS OF TRIGLYCINE UPON ADDITION OF Cu^{2+}

Copper Concentration (M)	T_{1P}^{-1} (sec^{-1})			T_{2P}^{-1} $(sec)^{-1}$			T_{1P}/T_{2P}		
	N	Central	C	N	Central	C	N	Central	C
1×10^{-5}	0.11	0.065	0.059	2.50	1.25	0.92	22.7	19.2	15.6
3.5×10^{-5}	0.33	0.25	0.26	4.00	2.03	1.35	12.1	8.1	5.2
2.4×10^{-4}	1.22	0.85	0.83	13.20	2.89	2.41	10.8	3.4	2.9
6.3×10^{-4}	2.22	2.12	2.08	19.60	5.51	4.58	8.8	2.6	2.2
12.0×10^{-4}	21.30	17.80	15.00	71.40	35.61	28.52	3.4	2.0	1.9

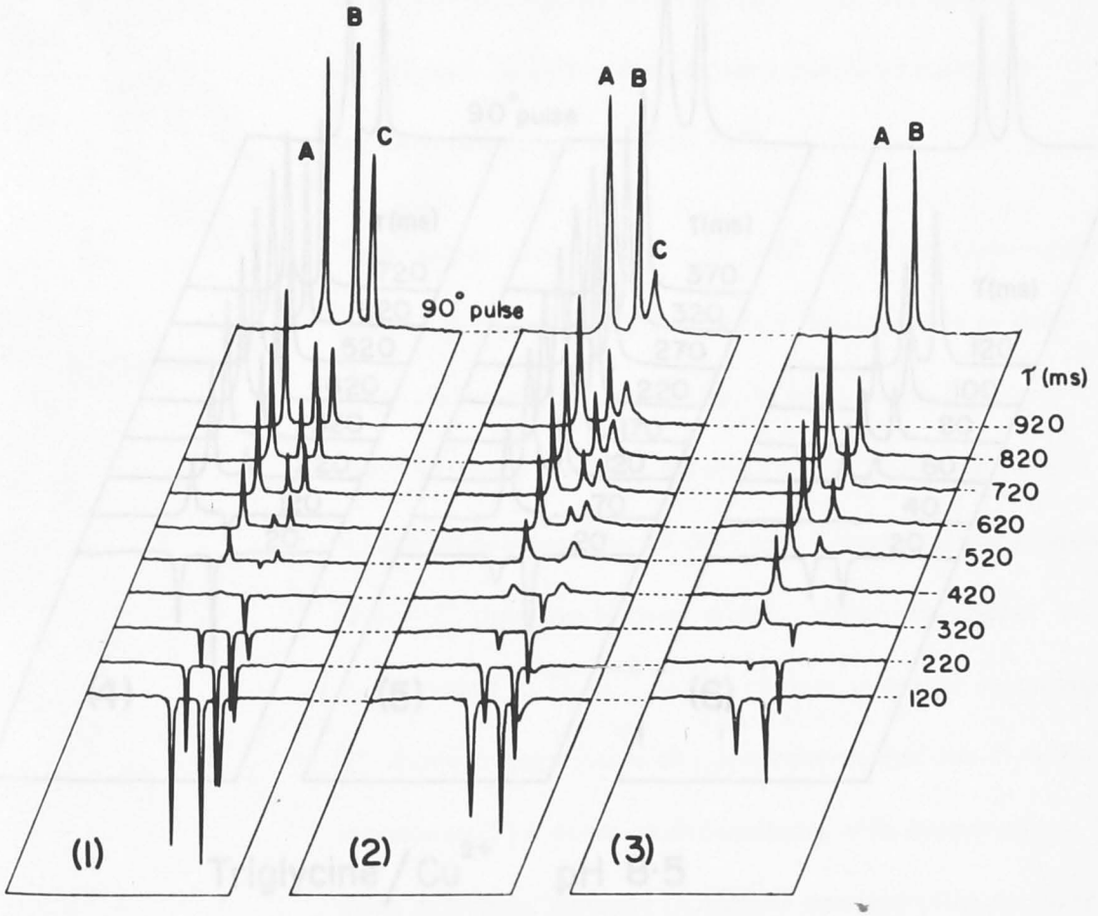


Fig. 5.4 Inversion recovery T_1 measurement, $(180 - \tau - 90)_n$ pulse sequence, for α-proton resonances in the presence of increasing Cu^{2+} concentrations of (1) Nil (2) 1×10^{-5}M. (3) 3.5×10^{-5}M. Resonances are (A) central proton (B) C-terminal α-proton, (C) N-terminal α-proton.

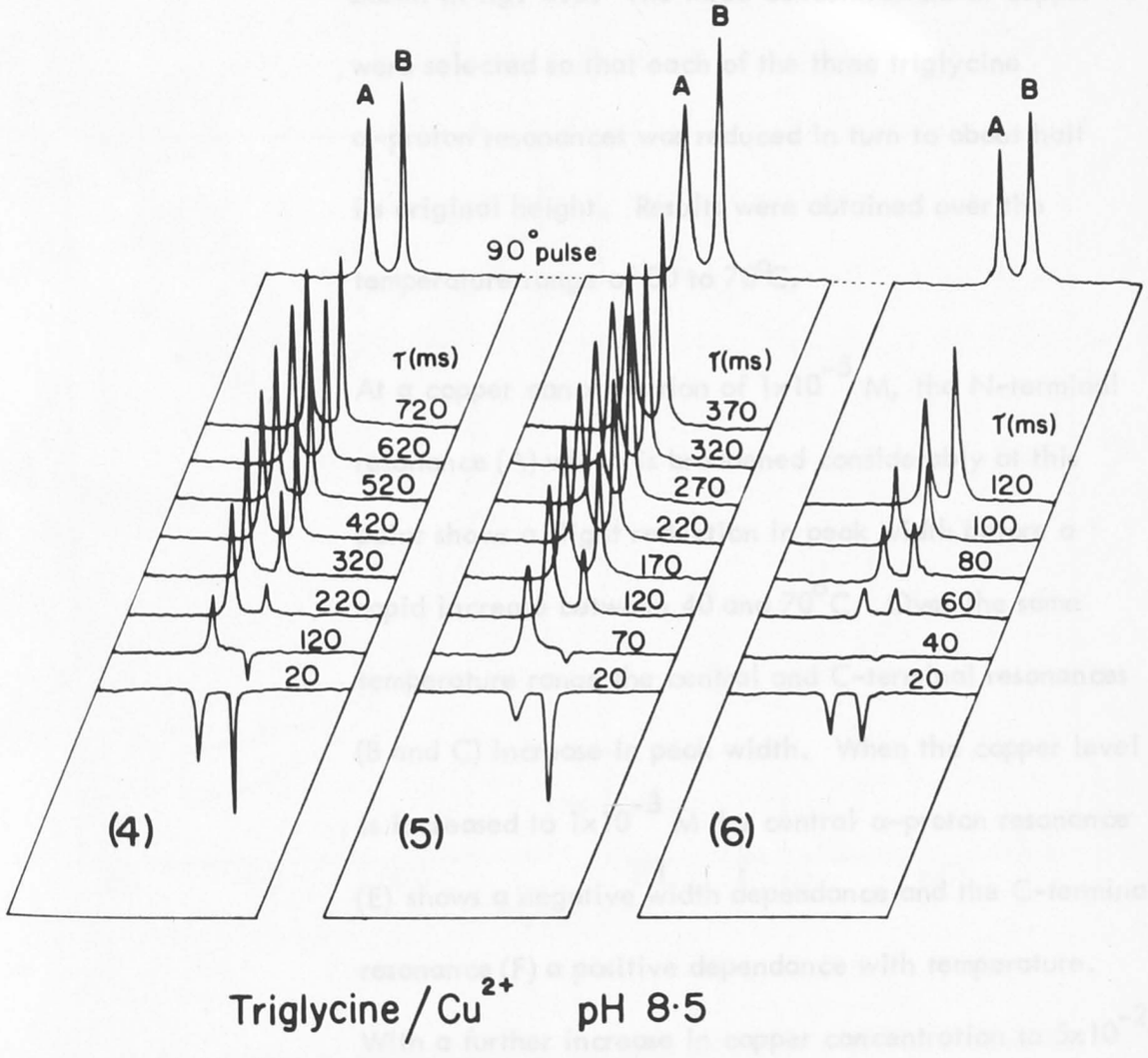


Fig. 5.4 (continued)

Copper concentrations are (4) 2.4×10^{-4}M. (5) 6.3×10^{-4}M. (6) 12.0×10^{-4}M.

C. Effect of temperature on broadening rates.

The effect of increasing temperature on the width at half height (Hz) of α-proton resonances of triglycine is shown in fig. 5.5. The three concentrations of copper were selected so that each of the three triglycine α-proton resonances was reduced in turn to about half its original height. Results were obtained over the temperature range of 30 to 70°C.

At a copper concentration of 1×10^{-5} M, the N-terminal resonance (A) which is broadened considerably at this point shows a slight reduction in peak width before a rapid increase between 40 and 70°C. Over the same temperature range the central and C-terminal resonances (B and C) increase in peak width. When the copper level is increased to 1×10^{-3} M the central α-proton resonance (E) shows a negative width dependance and the C-terminal resonance (F) a positive dependance with temperature. With a further increase in copper concentration to 5×10^{-2} M the remaining C-terminal α-proton resonance (G) now shows a negative width dependance with temperature.

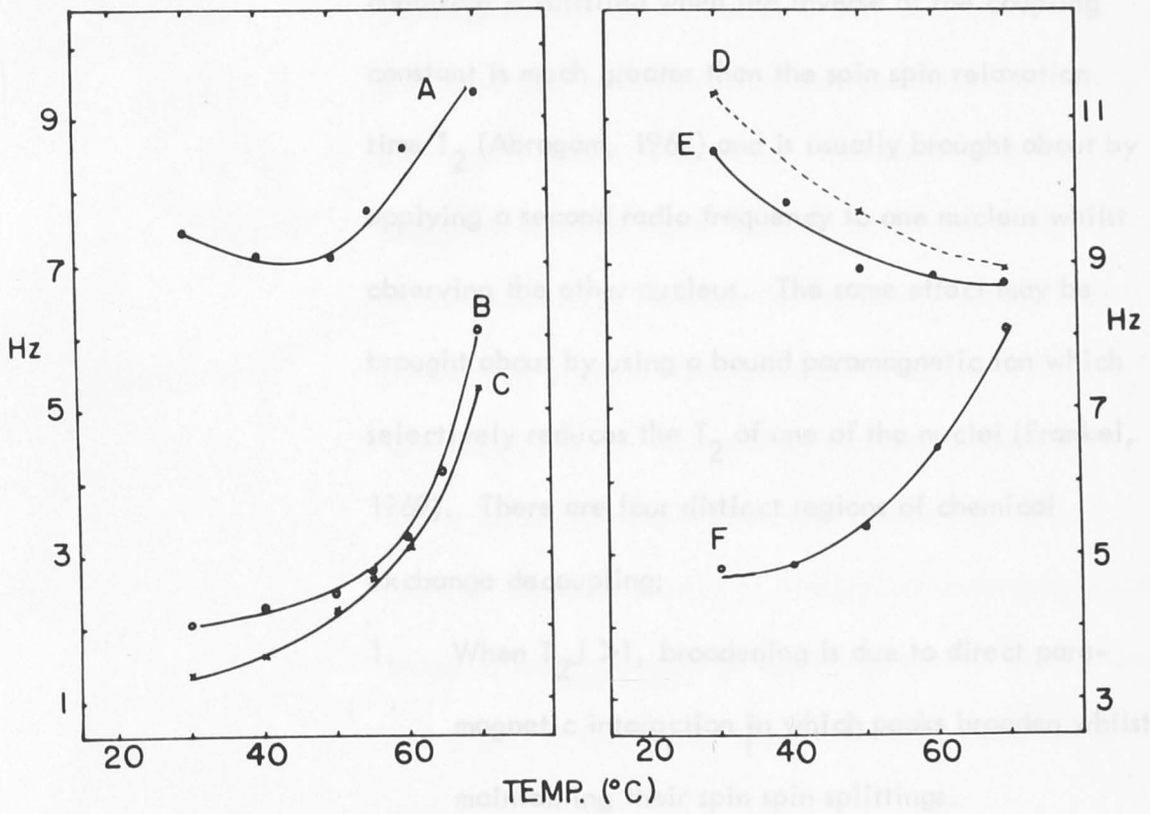


Fig. 5.5 Effect of temperature on the peak width of α-proton resonances of the copper triglycine complex in D_2O solution at pH 9.0. Peak widths are measured in Hz. at half peak height.

A. N-terminal ^1H resonance)
B. Central " ") Cu^{2+} concentration = 1×10^{-5}M.
C. C-terminal " ")

E. Central " ")
F. C-terminal " ") Cu^{2+} concentration = 1×10^{-4}M.

D. C-terminal " ") Cu^{2+} concentration = 1×10^{-2}M.

D. Chemical exchange decoupling.

In order to achieve spin spin decoupling the mean lifetime of a nucleus (A) in a certain spin state must be sufficiently shortened so that the coupled nucleus (B) can only sense the average field associated with nucleus (A). This condition is fulfilled when the inverse of the coupling constant is much greater than the spin spin relaxation time T_2 (Abragam, 1961) and is usually brought about by applying a second radio frequency to one nucleus whilst observing the other nucleus. The same effect may be brought about by using a bound paramagnetic ion which selectively reduces the T_2 of one of the nuclei (Frankel, 1969). There are four distinct regions of chemical exchange decoupling;

1. When $T_2J > 1$, broadening is due to direct paramagnetic interaction in which peaks broaden whilst maintaining their spin spin splittings.
2. At a point when $\sqrt{2} < T_2J < 5$, the separation of peaks just starts to decrease.
3. Then $0.5 < T_2J < \sqrt{2}$ the multiplet has collapsed to a singlet but T_2 is not rapid enough to produce a sharp line.
4. When $T_2J \ll 1$ complete spin spin decoupling occurs.

This effect was not noticed by Espersen (1976), who made a point of choosing ligands which did not have coupled nuclei such as triglycine. However, chemical exchange can give some insight into the relaxation mechanisms and therefore this effect has been studied in a number of peptides which show this effect. The compounds were chosen so that coupled nuclei were present at increasing distances from the N-terminus. The resultant effect of copper on each multiplet resonance is shown in fig. 5.6. As the distance of the observed nucleus increases from the N-terminus a higher concentration of copper is required to broaden the multiplet and an increase in chemical exchange decoupling is observed. Note that the alanine α-proton quartet in AlaGlyGly undergoes broadening without coalescence whilst the quartet in GlyAla and triplet in GlyGlySer show increasing coalescence with broadening.

These effects agree in a qualitative manner with the T_1/T_2 ratios for triglycine nuclei in the presence of copper, and can be explained by the presence of a contact interaction, a through-bond effect which decreases rapidly with distance from paramagnetic site. Hence α-protons nearest this site are broadened by this through-bond contact interaction which does not extend far enough

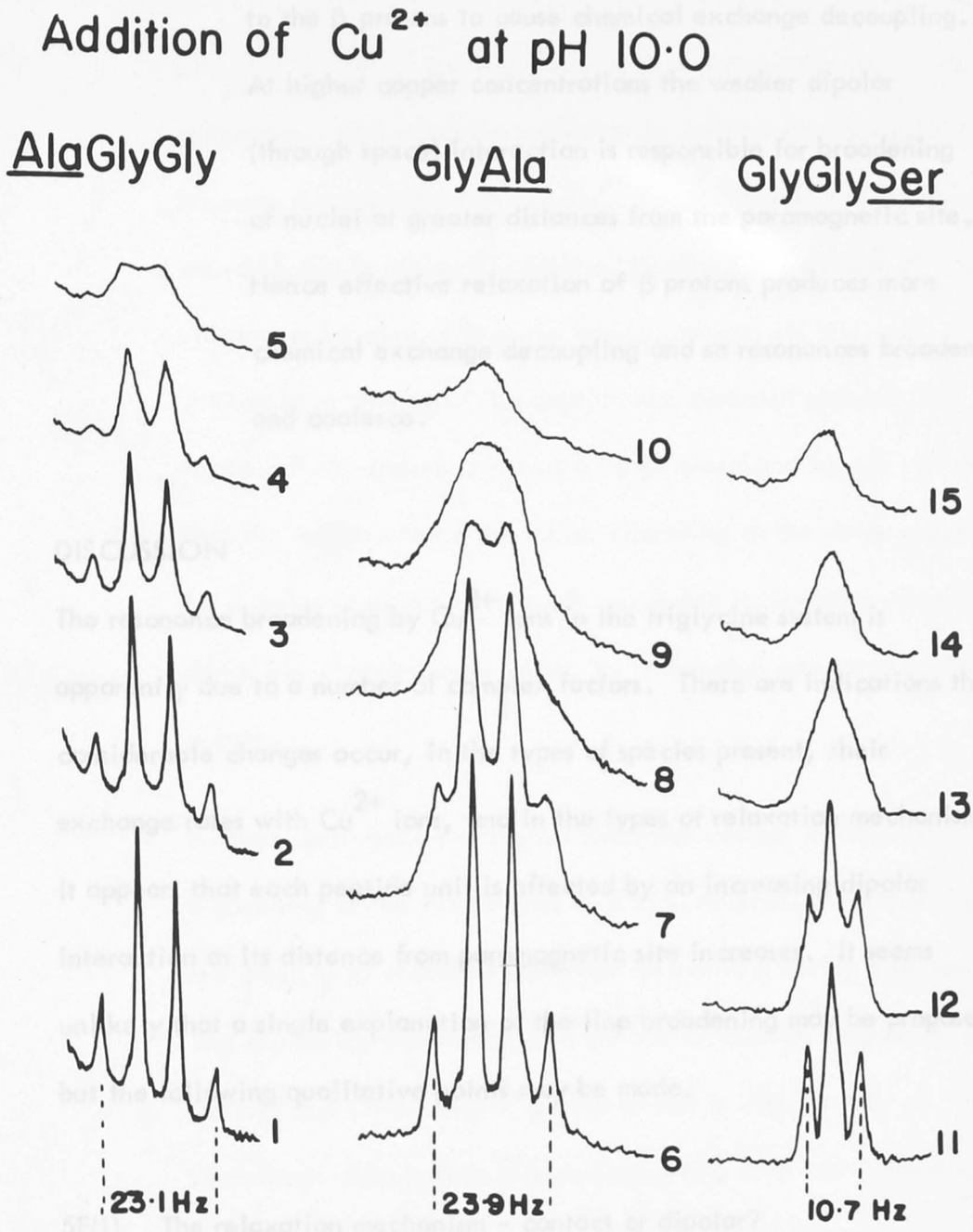


Fig. 5.6 Chemical exchange decoupling. The multiplet resonances for the α-protons of each amino acid residue (underlined) are shown over increasing copper concentration ranges of:

AlaGlyGly	1×10^{-6} to 1×10^{-5}M. (spectra 1 to 5).
GlyAla	1×10^{-5} to 1×10^{-3}M. (spectra 6 to 10).
GlyGlySer	1×10^{-4} to 1×10^{-2}M. (spectra 11 to 15).

to the β protons to cause chemical exchange decoupling. At higher copper concentrations the weaker dipolar (through space) interaction is responsible for broadening of nuclei at greater distances from the paramagnetic site. Hence effective relaxation of β protons produces more chemical exchange decoupling and so resonances broaden and coalesce.

5E. DISCUSSION

The resonance broadening by Cu^{2+} ions in the triglycine system is apparently due to a number of complex factors. There are indications that considerable changes occur, in the types of species present, their exchange rates with Cu^{2+} ions, and in the types of relaxation mechanisms. It appears that each peptide unit is affected by an increasing dipolar interaction as its distance from paramagnetic site increases. It seems unlikely that a single explanation of the line broadening may be proposed, but the following qualitative points may be made.

5E(i) The relaxation mechanism - contact or dipolar?

Measurements of T_{1p} / T_{2p} ratios show that for at least the N-terminal α-proton of triglycine, relaxation occurs via a contact or scalar interaction which is not distance related and therefore would not allow any structural information to be obtained. A dipolar interaction however small, might also be expected. In fact this dipolar through space, distance related mechanism appears to become more important for relaxation of the Central

and C-terminal α-proton nuclei, as evidenced by the reduction of T_{1p} / T_{2p} ratios towards the theoretical value for purely dipolar relaxation. This apparent reduction in contact interaction becomes more apparent as the Cu^{2+} concentration is increased and could be due to changes in species distribution from one of unidentate binding of terminal amino group to di, tri, and tetradentate binding of the peptide and carboxyl groups. In support of this theory is the note by Espersen and Martin (1976) that the scalar interaction varies according to the donor group and for a series of model compounds was shown to decrease when the donor was changed from amino, to π bonding amines to carboxyl groups.

Further confirmation of the predominent contact interaction for broadening of the N-terminal α-proton resonance is indicated by the absence of chemical exchange decoupling during broadening. The increasing inportance of a dipolar relaxation mechanism for Central and C-terminal α-proton nuclei is indicated by an increase in chemical exchange decoupling of α and β resonances as their distance from the paramagnetic site increases.

5E(ii) Exchange rates

Because there are many processes and correlation times which are temperature dependant, the affect of temperature change on broadening of resonances is difficult to interpret. In general, the triglycine system shows both broadening and narrowing as temperature is varied. If the exchange rate of Cu^{2+} were slow on the NMR time scale, then broadening would be governed by rate of chemical exchange and would be expected to increase with increasing temperature and hence cause increased broadening with temperature. This is not the case for N-terminal α-proton broadening and it seems more likely that in this case, where only trace levels of Cu^{2+} are required to completely broaden the resonance, the species responsible is in rapid exchange with Cu^{2+}. In fact only small amounts of an N-terminal bound complex of the type CuG_2, CuG_3, or CuG_4 would be required. At higher levels of Cu^{2+} where such species would be unlikely to exist broadening may be due to (a) a distance related relaxation mechanism with fast exchanging N-terminal bound species or (b) exchange broadening, resulting from di, tri, and tetradentate species which are known to exist at higher Cu^{2+} concentrations and between pH 7 and 10 form relatively slowly because of the rate limiting amide deprotonation step. In view of the temperature studies at these Cu^{2+} concentrations which show increased broadening with temperature the latter explanation seems more probable.

It is just conceivable that a 2nd order exchange process be responsible for broadening with species of type S (iii) showing slow chelation of one ligand with concomittant fast exchange of a further N-terminal bound ligand with bulk ligands.

5E(iii) Broadening at various pH values.

This study was undertaken to assess the conditions most suitable for peptide sequence studies. The variation in resonance broadening with pH shows that sequential broadening generally occurs over the range pH 7 to 10. At higher values this effect is gradually replaced by equal broadening of each α-proton resonance. The various regions examined were:

pH 1.0 This region showed virtually no copper binding, but just slight general broadening from the presence of free paramagnetic ions. At this pH the carboxyl group would be protonated and unlikely to bind copper ions.

pH 4.0 In this region there appears to be competition between the N-terminal group and the partially deprotonated carboxyl group for Cu^{2+} ions, resulting in broadening of the two terminal α-proton resonances before the central α-proton resonance.

pH 7 to 10 Sequential broadening from the N to C-terminus in this pH range is most likely caused by:

(1) Broadening of the N-terminal α-proton resonance at very low Cu^{2+} concentrations from an N-terminal bound species undergoing fast exchange with the bulk ligands.

(2) Sequential broadening of central and C-terminal α-proton resonances caused by slow exchange processes in formation of di, tri, and tetradentate complexes.

pH 10 to 13 The change to equal broadening of all resonances in triglycine copper complexes could be explained by an increase in the rate of amide deprotonation with increase in pH. pH 10 appears to be close to the unknown value at which the experiments of Kim and Martell (1969) were carried out. This appears to be a value where exchange broadening starts to become rapid on the NMR time scale causing equal broadening of central and C-terminal resonances.

5F. APPLICATION TO PEPTIDE SEQUENCING

Previous measurements of T_1 and T_2 for protons of triglycine have indicated that T_1 values are unlikely to be of any use for peptide sequencing but that T_2 values (peak broadening) could be used as a basis for sequencing at least tripeptides from the N-terminus, in the pH range 7 to 10. Peptides may be classified into three groups; those with neutral, acidic or basis side chains.

5F(i) Peptides with neutral side chains.

In this group Cu^{2+} has proved a useful probe for sequence determination of up to 4 residues. Polyglycines are the most simple group of peptides to sequence and tetraglycine the longest one in which it has been possible to demonstrate complete sequence (fig. 5.7). It is not possible to say with certainty whether this is the maximum length because the limitation is one of spectral resolution which deteriorates rapidly as the broadened resonances merge together. When other neutral amino acid residues are present spin spin splitting results in proton resonances of doublets through to quadruplet peaks. Hence apparent resolution is decreased and also the area of each α-proton multiplet is half that of the α-proton glycine resonance. It is still possible to sequence some of these peptides but extreme caution is essential in interpretation. One recurring problem is that of chemical exchange spin decoupling (Frankel, 1969) which results in the collapse of multiplets prior to or at the same time as broadening. The two effects may initially cancel out and the resonance may not appear to be broadening at all, and can be seen in the sequence determination of AlaLeuGly, and AlaGlyGly. (fig. 5.8 & 5.9). Other peptides in this group which have been successfully sequenced with copper include: IleGlyGly, GlyGlySer, and GlyLeuGlyLeu.

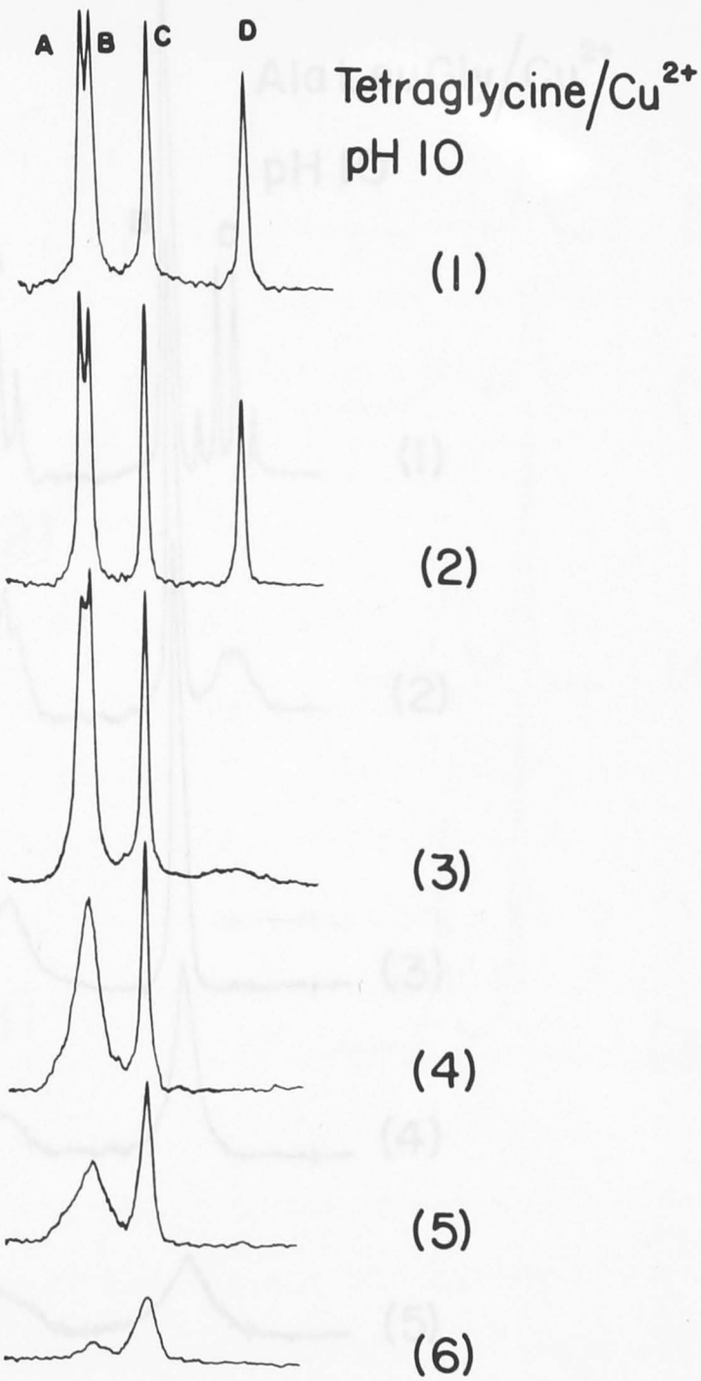


Fig. 5.7 PMR spectra of the α-proton region of tetraglycine showing the effect of added Cu^{2+} at molar concentrations of (1) Nil, (2) 0.5×10^{-5} (3) 0.5×10^{-4} (4) 5×10^{-4} (5) 1×10^{-3} (6) 5×10^{-3}, to a 0.1M solution of tetraglycine in D_2O. Peaks in sequence from the N-terminus are labelled D, A, B, C.

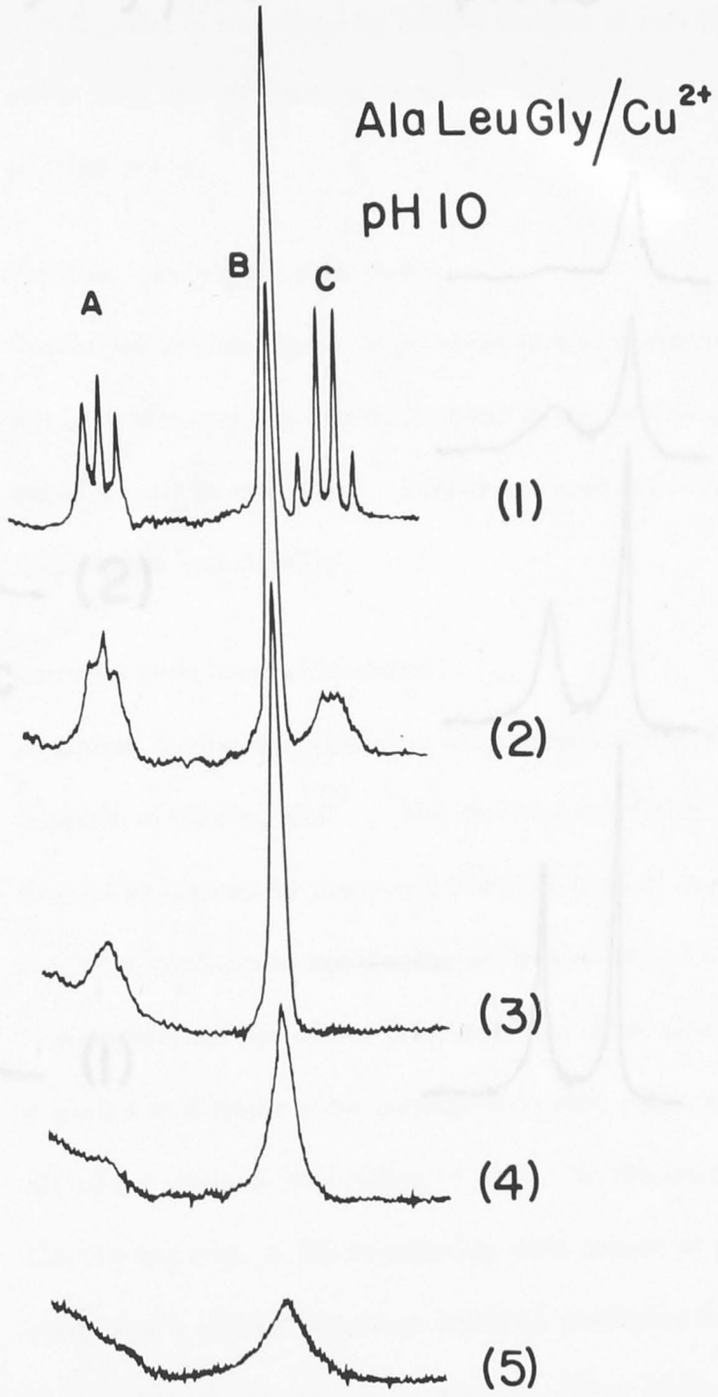


Fig. 5.8 PMR spectra of the α-proton region of AlaLeuGly showing the effect of increasing molar Cu^{2+} concentrations of (1) Nil, (2) 0.5×10^{-5}, (3) 1×10^{-4}, (4) 1×10^{-3}, (5) 5×10^{-3} the α-proton resonances are labelled as follows:

A. Triplet of leucine, B. Glycine singlet, C. Alanine quartet.

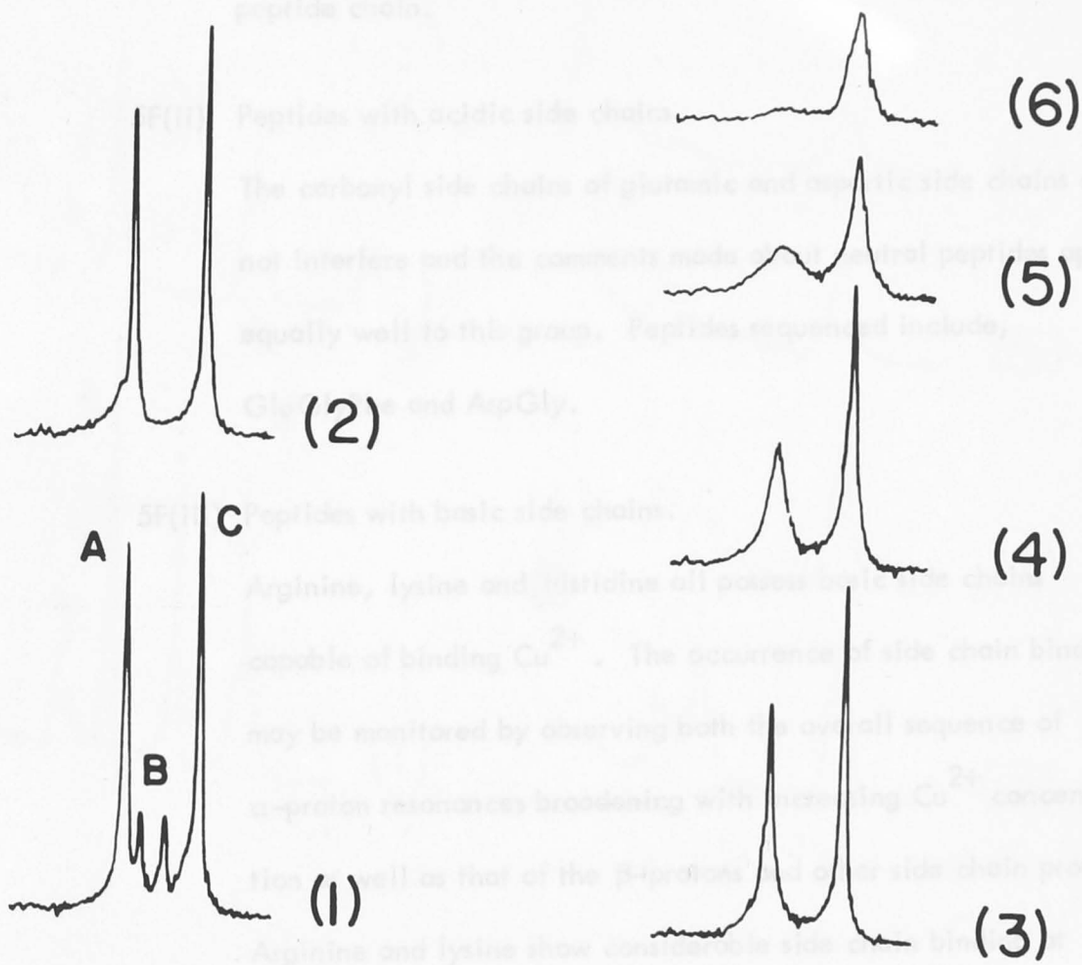


Fig. 5.9 PMR spectra of the α-proton region of AlaGlyGly with increasing molar concentrations of copper (Cu^{2+}) of (1) Nil, (2) 1×10^{-5}, (3) 5×10^{-5}, (4) 1×10^{-4}, (5) 1×10^{-3}, (6) 5×10^{-3}. Assignments are (A) Central glycine α-proton, (B) Overlapping quartet of alanine α-proton (C) C-terminal glycine α-proton.

In general, the difficulty of sequence determination of peptides in this group is increased by (a) the number of non-glycine amino acids units and (b) their nearness to the C-terminal group in the peptide chain.

5F(ii) Peptides with acidic side chains.

The carboxyl side chains of glutamic and aspartic side chains do not interfere and the comments made about neutral peptides apply equally well to this group. Peptides sequenced include, GluGlyPhe and AspGly.

5F(iii) Peptides with basic side chains.

Arginine, lysine and histidine all possess basic side chains capable of binding Cu^{2+}. The occurrence of side chain binding may be monitored by observing both the overall sequence of α-proton resonances broadening with increasing Cu^{2+} concentration as well as that of the β-protons and other side chain protons. Arginine and lysine show considerable side chain binding at pH 10 but little or no binding at pH 8. In the case of GlyGlyArg (fig. 5.10) broadening with copper at pH 8 it is clear that a correct sequence could be predicted because there is little broadening of the β-protons. Other basic peptides successfully sequenced at this pH include LysTrpLys, GlyLys, ArgProArg.

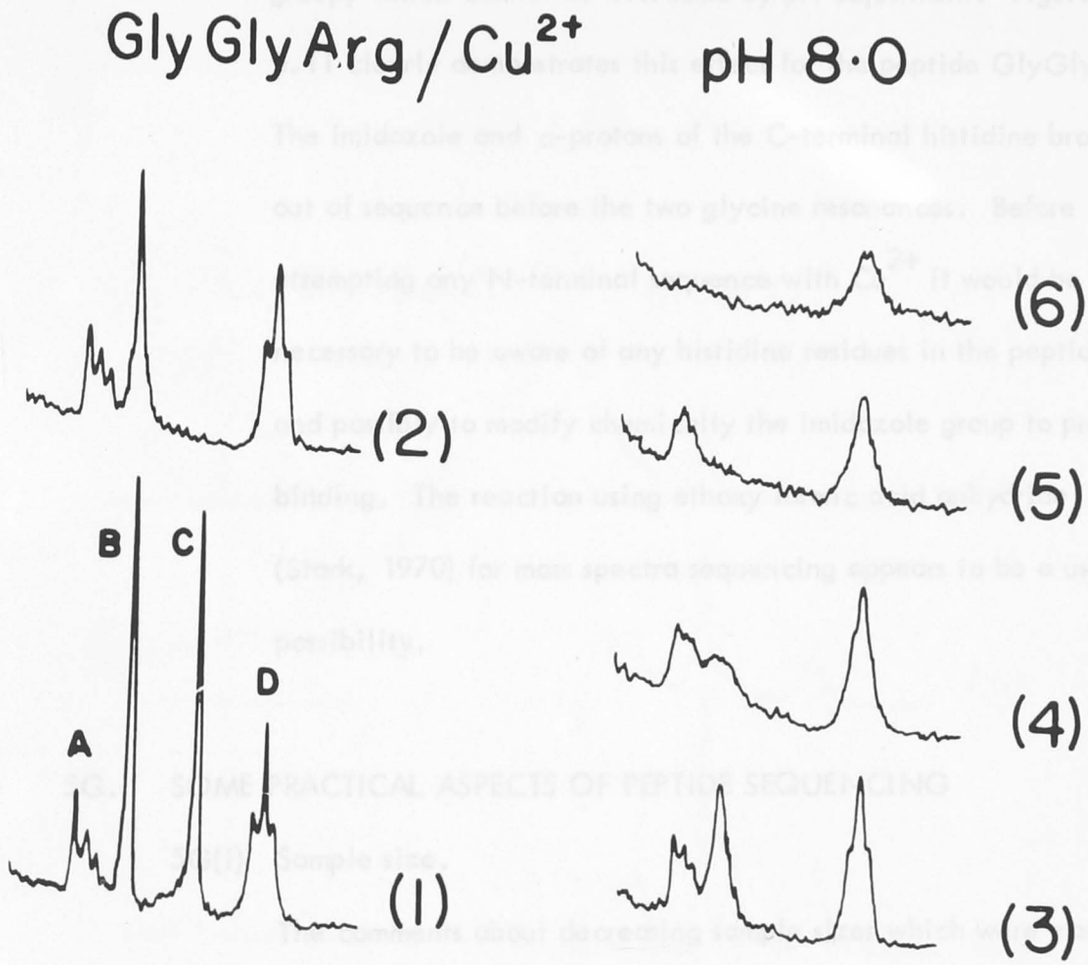


Fig. 5.10 NMR spectra of the α-proton region of GlyGlyArg with increasing molar concentrations of Cu^{2+} of (i) Nil, (2) 1×10^{-5}, (3) 1×10^{-4}, (4) 5×10^{-4}, (5) 1×10^{-3}, (6) 1×10^{-2}. Assignments are (A) Arg α ^1H (B) Central glycine α ^1H (C) C-terminal α ^1H (D) γ ^1H of arginine.

Histidine residues show considerable binding of Cu^{2+} to the imidazole group, which cannot be overcome by pH adjustment. Figure 5.11 clearly demonstrates this effect for the peptide GlyGlyHis. The imidazole and α-protons of the C-terminal histidine broaden out of sequence before the two glycine resonances. Before attempting any N-terminal sequence with Cu^{2+} it would be necessary to be aware of any histidine residues in the peptide, and possibly to modify chemically the imidazole group to prevent binding. The reaction using ethoxy formic acid anhydride (Stark, 1970) for mass spectra sequencing appears to be a useful possibility.

5G. SOME PRACTICAL ASPECTS OF PEPTIDE SEQUENCING

5G(i) Sample size.

The comments about decreasing sample sizes which were made in Chapter 4 equally apply in the case of copper complexes. The adjustable micro tube works well for copper complexes using sub milligram amounts of peptide. For example the sequential broadening of 25 mg. of AlaLeuGly in a standard NMR tube (fig. 5.10) can be repeated with 0.8 mg. in a volume of 20 μl in the micro tube (fig. 5.12). Both spectra were obtained with a continuous wave spectrometer and it should be possible to further reduce this amount to about 0.1 mg. using a pulse spectrometer.

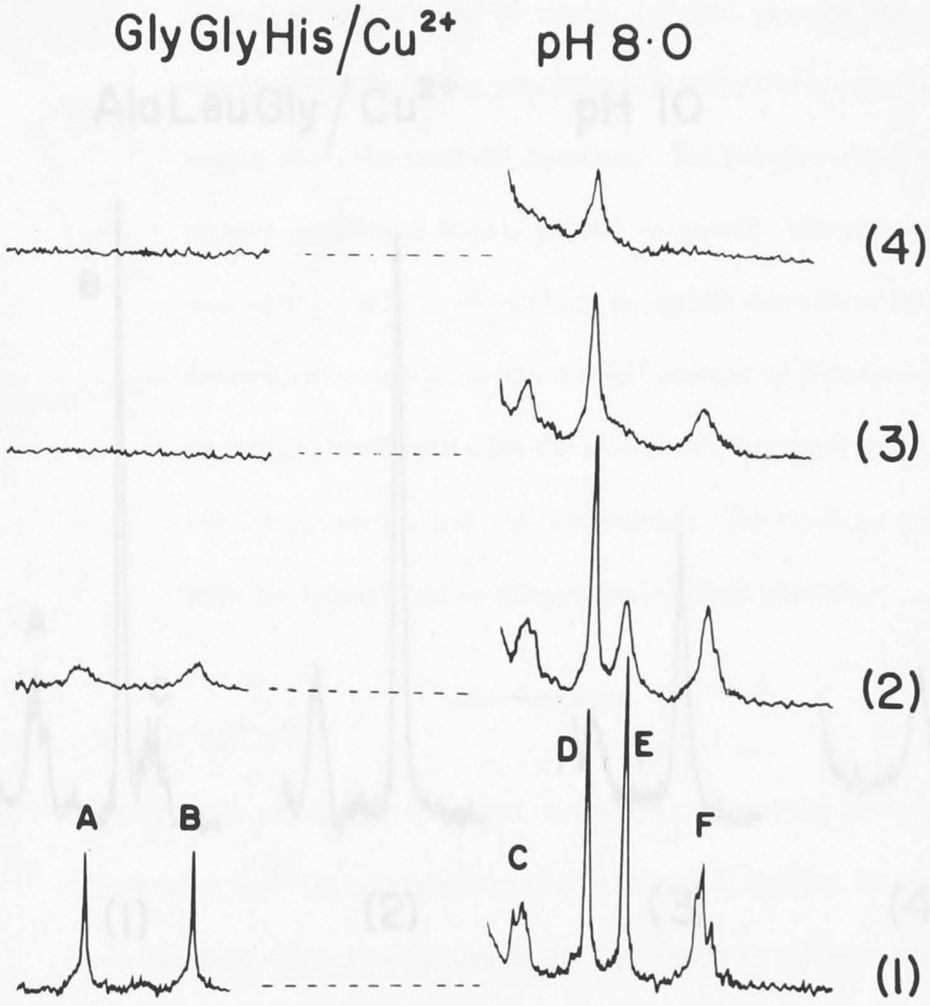


Fig. 5.11 PMR spectra of proton resonances of GlyGlyHis, in the presence of increasing Cu^{2+} molar concentrations of (1) Nil, (2) 1×10^{-5}, (3) 1×10^{-4}, (4) 1×10^{-3}. Assignments are (A) Imidazole C2 proton (B) Imidazole C4 proton (C) Histidine α ^1H (D) Central α ^1H of glycine (E) N-terminal α^1H (H) Histidine β ^1H.

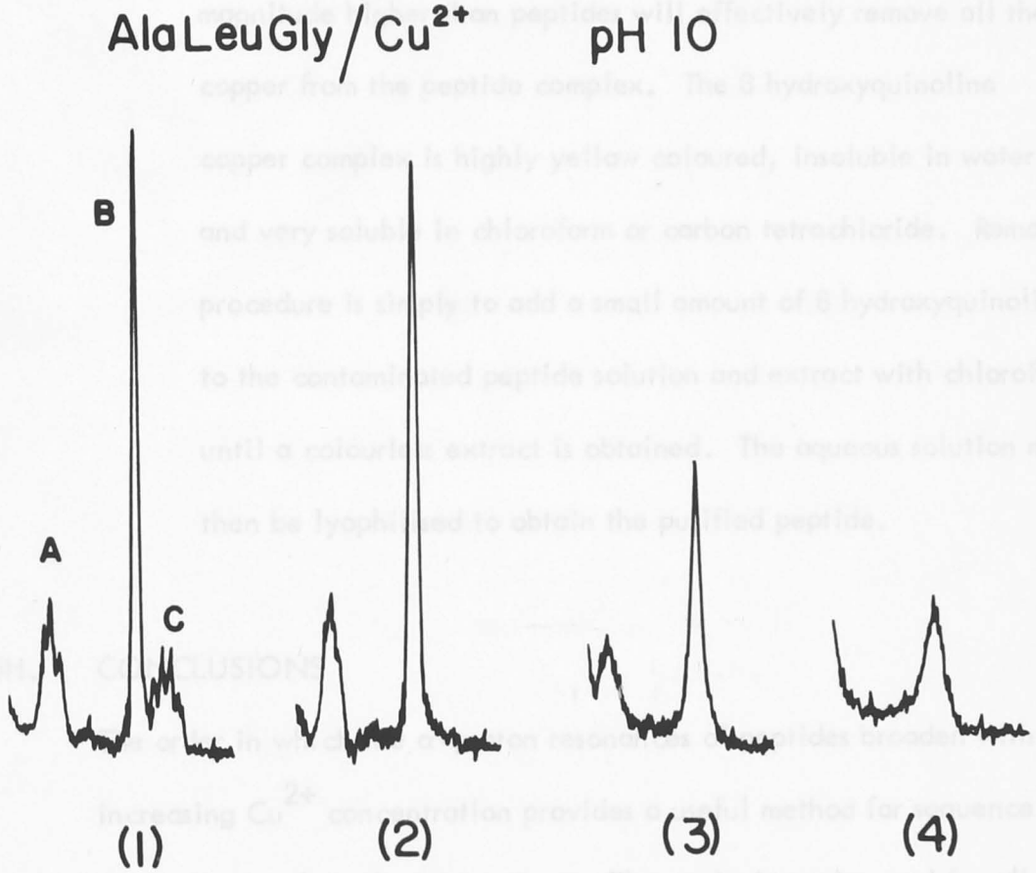


Fig. 5.12 100 MHz PMR spectra of 0.8 mg. of AlaLeuGly in 20 µl of D_2O contained in an adjustable micro bulb NMR tube, and showing the effect of increasing molar concentrations of Cu^{2+} of (1) Nil, (2) 1×10^{-5}, (3) 1×10^{-3}, (4) 5×10^{-3}. Assignments are (A) leucine triplet (B) glycine singlet (C) alanine quartet.

5G(ii) Recovery of peptides.

Sequencing by this method results in contamination of the peptide with Cu^{2+} ions. The peptide may be conveniently recovered free from Cu^{2+} by complexing the copper with a water insoluble strongly chelating ligand. 8 hydroxyquinoline which has a formation constant for its copper complex many orders of magnitude higher than peptides will effectively remove all the copper from the peptide complex. The 8 hydroxyquinoline copper complex is highly yellow coloured, insoluble in water and very soluble in chloroform or carbon tetrachloride. Removal procedure is simply to add a small amount of 8 hydroxyquinoline to the contaminated peptide solution and extract with chloroform until a colourless extract is obtained. The aqueous solution may then be lyophilised to obtain the purified peptide.

5H. CONCLUSIONS

The order in which the α-proton resonances of peptides broaden with increasing Cu^{2+} concentration provides a useful method for sequence determination from the N-terminus. The method may be used for all amino acid residues except histidine although an N-terminal histidine residue should not interfere and might well be beneficial in improving the sequencing. The length of peptide which may be sequenced by this method is at least four amino acid units except in cases where resonances from each α-proton overlap. The sequence determination can be adapted to sub milligram amounts of peptides using micro NMR tubes.

For example the sequence of 0.8 mg of AlaLeuGly using a single scan on a continuous wave spectrometer has been demonstrated and this quantity could be reduced even further with a FT pulse spectrometer, and multiple scans.

CHAPTER 6

THE USE OF ^{13}C NMR FOR PEPTIDE SEQUENCE DETERMINATION

6A. INTRODUCTION

Natural abundance of the ^{13}C nucleus is only 1.1% of total carbon nuclei, and at the same magnetic field strength its sensitivity is only 1.6% that of ^{1}H for equal numbers of nuclei. Hence the overall relative sensitivity of ^{13}C NMR is only 1.8×10^{-4} that of ^{1}H NMR. Observation of ^{13}C nuclei with continuous wave spectrometers is not possible and some form of signal enhancement is essential. This may include:

(1) Proton noise decoupling to remove all ^{1}H - ^{13}C spin spin splittings, resulting in singlet ^{13}C resonances.

(2) Computer accumulation of data from multiple scans. In this case signal/noise ratio increase as the square root of the number of scans.

(3) Fourier transform (FT) pulse techniques.

(4) ^{13}C enrichment of samples.

The lack of sensitivity represents the biggest single disadvantage of ^{13}C over ^{1}H NMR, but there are some important advantages to be gaining by using CMR which include:

(1) The range of chemical shifts for ^{13}C nuclear resonances is very much greater than for ^{1}H. For example the α-carbon resonances of peptides occupy a range of some 11 ppm whilst the α-proton resonances are spread only over 1.5 ppm. This results in less overlap of ^{13}C resonance from different amino acid residues.

(2) Spectra are first order, i.e. no ^{13}C - ^{13}C splittings.

(3) There is no need to use deuterated solvents and hence no problems occur from overlap of resonances with a solvent peak.

The superior resolution and the recent advances in instrumentation have made ^{13}C NMR a very attractive technique. An example of the resolution one might expect with peptides was recently published by Freedman et al, (1973) who was able to resolve and assign all the αC resonances from a peptide composed of the first 20 residues of ribonuclease A, a task that would certainly prove impossible with PMR even using the highest field spectrometers available.

A further question about the use of relaxation probes with ^{13}C NMR is the distance over which the effect of paramagnetic ions might be useful. The equations of Solomon (1955) and Bloembergen (1957) predict that the dipolar contribution to relaxation is proportional to the square of the gyro magnetic ratio and is independant of the Larmour precessional frequency provided that there is rapid motion as occurs with small molecules. The value for ^{13}C nuclei is approximately ¼ times that for ^{1}H nuclei. Therefore one might expect to need 16 times ar much

paramagnetic ion to broaden ^{13}C resonances to the same extent as ^{1}H resonances. Inevitably this would lead to the presence of more free Gd^{3+} ions in solution resulting in more general broadening and a shorter effective distance of sequential relaxation.

6B. RELATIVE BROADENING EFFECTS OF Cu^{2+} AND Gd^{3+} ON ^{1}H AND ^{13}C RESONANCES

To study the above mentioned effect, measurements of broadening of the triglycine α-proton resonances at pH 4.0 as a function of concentration of Gd^{3+} were made for both ^{13}C and ^{1}H nuclei. (fig. 6.1). Similar experiments were also carried out to compare the effect of added Cu^{2+} at pH 8.0 with PMR and CMR. Many more spectra were recorded than those actually shown and fractional peak heights were graphed against added concentrations of paramagnetic ion. The concentration ratios were determined for a 50% reduction in peak height and gave values of:

(1) $\lceil Cu^{2+} \rfloor_{CMR} / \lceil Cu^{2+} \rfloor_{PMR} = 3 \pm 2$

(2) $\lceil Gd^{3+} \rfloor_{CMR} / \lceil Gd^{3+} \rfloor_{PMR} = 12 \pm 5$

The high degree of variability is caused by noise levels in the CMR spectra and in the variability of heights in ^{13}C resonances resulting from changes in NOE and relaxation rates in the presence of paramagnetic ions. However, the value of 12 for Gd^{3+} is close to the theoretical value of 16. In the case of Cu^{2+}, the contact term is important and it is not possible to predict the result from theory.

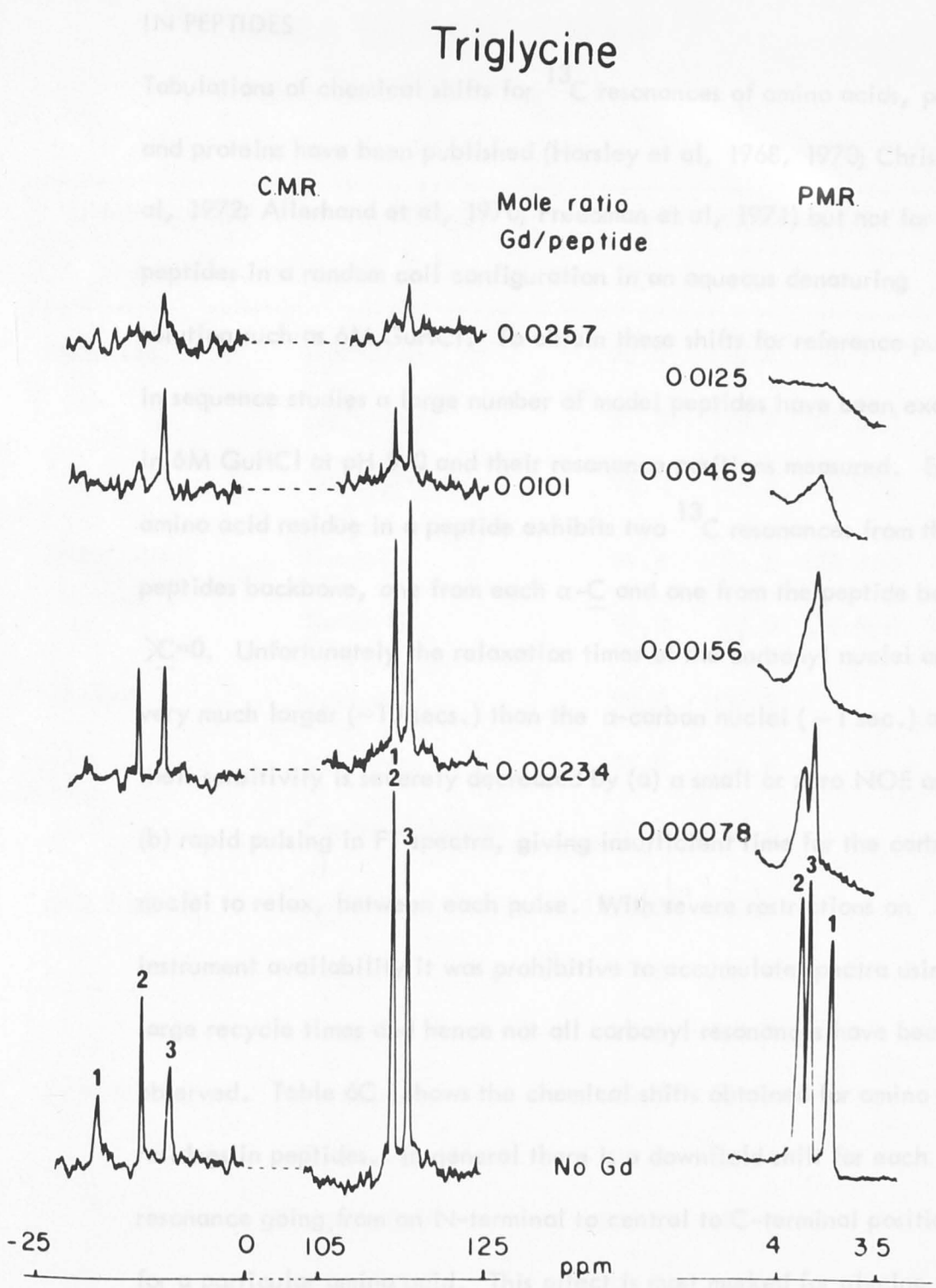


Fig. 6.1 Comparison of the amounts of Gd^{3+} required to broaden PMR and CMR resonances of Triglycine. Peak numbers refer to nuclei numbered sequentially from the N-terminus.

6C. CHEMICAL SHIFTS OF ^{13}C MAGNETIC NUCLEAR RESONANCES IN PEPTIDES

Tabulations of chemical shifts for ^{13}C resonances of amino acids, peptides and proteins have been published (Horsley et al, 1968, 1970; Christle et al, 1972; Allerhand et al, 1970; Freedman et al, 1971) but not for peptides in a random coil configuration in an aqueous denaturing solution such as 6M GuHCl. To obtain these shifts for reference purposes in sequence studies a large number of model peptides have been examined in 6M GuHCl at pH 5.0 and their resonance positions measured. Each amino acid residue in a peptide exhibits two ^{13}C resonances from the peptides backbone, one from each α-$\underline{C}$ and one from the peptide bond $\underline{C}$=0. Unfortunately the relaxation times of the carbonyl nuclei are very much larger (~10 secs.) than the α-carbon nuclei (~1 sec.) and their sensitivity is severely decreased by (a) a small or zero NOE and (b) rapid pulsing in FT spectra, giving insufficient time for the carbonyl nuclei to relax, between each pulse. With severe restrictions on instrument availability it was prohibitive to accumulate spectra using large recycle times and hence not all carbonyl resonances have been observed. Table 6C I shows the chemical shifts obtained for amino acid residues in peptides. In general there is a downfield shift for each resonance going from an N-terminal to central to C-terminal position for a particular amino acid. This affect is most marked for glycine and reversed for the central resonance of serine. The variability of peak positions resulting from different adjacent amino acid residues indicated by glycine containing peptides (Table 6C II) is shown to average 0.3 ppm

TABLE 6C I

^{13}C CHEMICAL SHIFTS FOR PEPTIDES IN

6M GUANIDINE HCl AT pH 4.0

AMINO ACID	PEPTIDE	SHIFT RELATIVE TO GuHCl			
		αC	βC		>C=0
PROLINE	ProPheAsp	98.0	128.4	110.0	*
	ArgProArg	97.2	128.4	109.6(S)	*
	GlyGlyPro	96.0	128.5	110.9(S)	*
THREONINE	ThrLysTyr	99.4	91.5	138.8(γ)	- 9.9
	ThrGlyGly	99.0	91.5	138.8(γ)	-10.9
VALINE	ValTyrVal	99.5	127.8		*
	ValTyrVal	96.9	127.4		*
	PheAspAlaSerVal	96.8	127.7		*
ISOLEUCINE	IleIleIle	100.2	122.1		-10.6
	IleIleIle	99.8	121.9		-14.3
	IleIleIle	98.3	121.7		-19.4
TYROSINE	TyrGlyGly	103.0	121.8		-12.0
	ValTyrVal	102.5	121.4		*
	GlyTyrGly	102.4	121.3		-14.9
	GlyGlyTyr	101.3	121.3		-20.3
TRYTOPHAN	TrpGlyGly	103.9	130.9		*
	LysTrpLys	102.9	131.4		-14.1
	GlyGlyTrp	102.1	130.5		-20.1

TABLE 6C I

(continued)

AMINO ACID	PEPTIDE	SHIFT RELATIVE TO GuHCl			
		αC	βC		C=0
SERINE	SerGlyGly	103.0	97.6		-10.4
	SerGlyAlaGlyAlaGly	102.0	95.5		*
	GlySerAla	102.4	96.5		-12.7
	PheAspAlaSerVal	102.3	96.8		-13.1
	GlyAspSer	100.7	95.9		-18.0
LYSINE	LysLysLys	105.0	127.3	118.2(ε)	-10.4
	LysTryLys	104.6	127.3	118.5(ε)	-11.0
	ThrLysTyr	103.9	127.1	118.2(ε)	-15.0
	LysLysLys	103.9	127.3	118.2(ε)	-14.5
	LysTryLys	102.9	127.1	118.5(ε)	-19.4
	LysLysLys	102.6	127.3	118.2(ε)	-19.2
ARGININE	ArgProArg	106.0	130.9	116.8(δ)	*
	ArgProArg	102.7	129.5	116.8(δ)	*
	GlyGlyArg	102.9	129.5	116.9(δ)	*
PHENYLA LANINE	PheGlyGly	103.4	121.3		-11.4
	PheAspAlaSerVal	103.6	121.2		-10.9
	ProPheAsp	102.0	121.4		*

TABLE 6C I

(continued)

AMINO ACID	PEPTIDE	SHIFT RELATIVE TO GuHCl			
		αC	βC		>C=0
GLUTAMIC ACID	GluGluGlu	105.2	131.1	126.5(γ)	-11.3
	GluGluGlu	104.4	131.1	126.5(γ)	-14.3
	GluGluGlu	103.4	131.1	126.5(γ)	-20.3
HISTIDINE	HisGlyGly	105.5	131.6		*
	GlyHisGly	105.1	130.8		-13.38
METHIONINE	MetMetMet	106.9	128.5		-11.3
	MetMetMet	105.2	127.5		-14.1
	MetMetMet	104.8	127.1		*
	GlyGlyMet	103.5	127.1		*
LEUCINE	LeuLeuLeu	105.8			*
	LeuLeuLeu	105.4			*
	LeuLeuLeu	103.8			*
GLUTAMINE	GlnGlnGln	106.9	132.8		*
	GlnGlnGln	106.2	132.8		*
	GlnGlnGln	104.9	132.8		*

TABLE 6C I

(continued)

AMINO ACID	PEPTIDE	SHIFT RELATIVE TO GuHCl		
		αC	βC	>C=0
ASPARTIC ACID	GlyAspSer	106.7	119.9	-14.3
	PheAspAlaSerVal	106.8	119.7	-13.6
	ProPheAsp	105.0	119.6	*
ALANINE	AlaGlyGly	108.4	141.2	-13.3
	PheAspAlaSerVal	108.2	141.2	-16.4
	SerGlyAlaGlyAlaGly	108.4/ 108.9	141.7/ 142.1	*
	GlySerAla	106.8	140.8	

Notes: Shifts are quoted in ppm upfield from guanidine hydrochloride (internal standard) which occurred 159.4 ppm downfield from external TMS. A negative value indicates a downfield shift from GuHCl.

* Resonance not observed because of weak solution (low solubility of peptide) and possibly longer T_1 values of these nuclei.

TABLE 6C II

^{13}C CHEMICAL SHIFTS FOR GLYCINE IN PEPTIDES

AT pH 4.0 IN GUANIDINE HYDROCHLORIDE

PEPTIDE	αC			>C = 0		
	N-terminal	Central	C-terminal	N-terminal	Central	N-terminal
GlyGlyPro	117.2	116.0	-	*	*	-
GlyGlyTyr	117.5	115.9	-	-9.9	-12.6	-
GlyGlyMet	116.9	115.2	-	*	*	-
GlyGlyTrp	117.2	115.2	-	-9.6	-12.4	-
GlyGlyArg	116.9	115.2	-	*	*	-
GlyAspSer	116.9	-	-	-9.2	-	-
GlySerAla	117.1	-	-	-9.5	-	-
GlyTyrGly	117.1	-	114.5	-8.8	-	-17.4
GlyHisGly	116.8	-	114.3	-9.2	-	-18.6
ThrGlyGly	-	115.0	114.2	-	-12.8	-18.8
SerGlyGly	-	115.1	114.5	-	-12.8	-18.3
TyrGlyGly	-	115.2	114.4	-	-12.7	-18.4
TrpGlyGly	-	115.1	114.5	-	*	*
PheGlyGly	-	115.4	114.6	-	-12.2	-17.8
HisGlyGly	-	115.3	114.6	-	*	*
AlaGlyGly	-	115.2	114.6	-	-12.8	-18.1

Notes: Shifts are quoted in ppm upfield from guanidine hydrochloride (internal standard) which occurred 159.4 ppm downfield from external TMS.

* Resonances not observed because of weak solution (low solubility of peptide) and possibly longer T_1 values of these nuclei.

for α-carbons and 0.9 ppm for carbonyl resonances (excluding carboxyl groups because of their variable degree of charging). The α-C resonances therefore appear to be more suitable for sequence studies because of the lesser relative effect of adjacent residues on their chemical shifts.

Resonances from terminal amino acid units also show pH shifts reflecting the protonation of carboxyl and amino groups. This effect for triglycine (fig. 6.2) results in an upfield shift of about 1.5 and 1.0 ppm during protonation of amino and carbonyl α-C nuclei respectively but is reversed for the carbonyl resonances and results in downfield shifts of 8 and 1.5 ppm respectively. On protonation of a particular group one would expect that adjacent hydrogens and carbons would be deshielded and show downfield chemical shifts. This is indeed the case for hydrogen nuclei but is reversed for ^{13}C resonances. Horsley and Sternlicht (1968) have explained this in terms of two opposing effects, (1) a through space electric field effect and (2) a through bond inductive effect which falls off more rapidly with distance. Such anomalous pH shifts were noted by Freedman (1971) for L Histidine and Gurd et al, (1971) for ValLeuSerGluGly. In both cases the α-C and carbonyl resonances showed shifts in the same direction. In the case of triglycine in 6M GuHCl solution however carbonyl resonances shift in the opposite direction to α-C resonances during protonation (fig. 6.2). Obviously with such variations, great care must be exercised in assigning ^{13}C resonances.

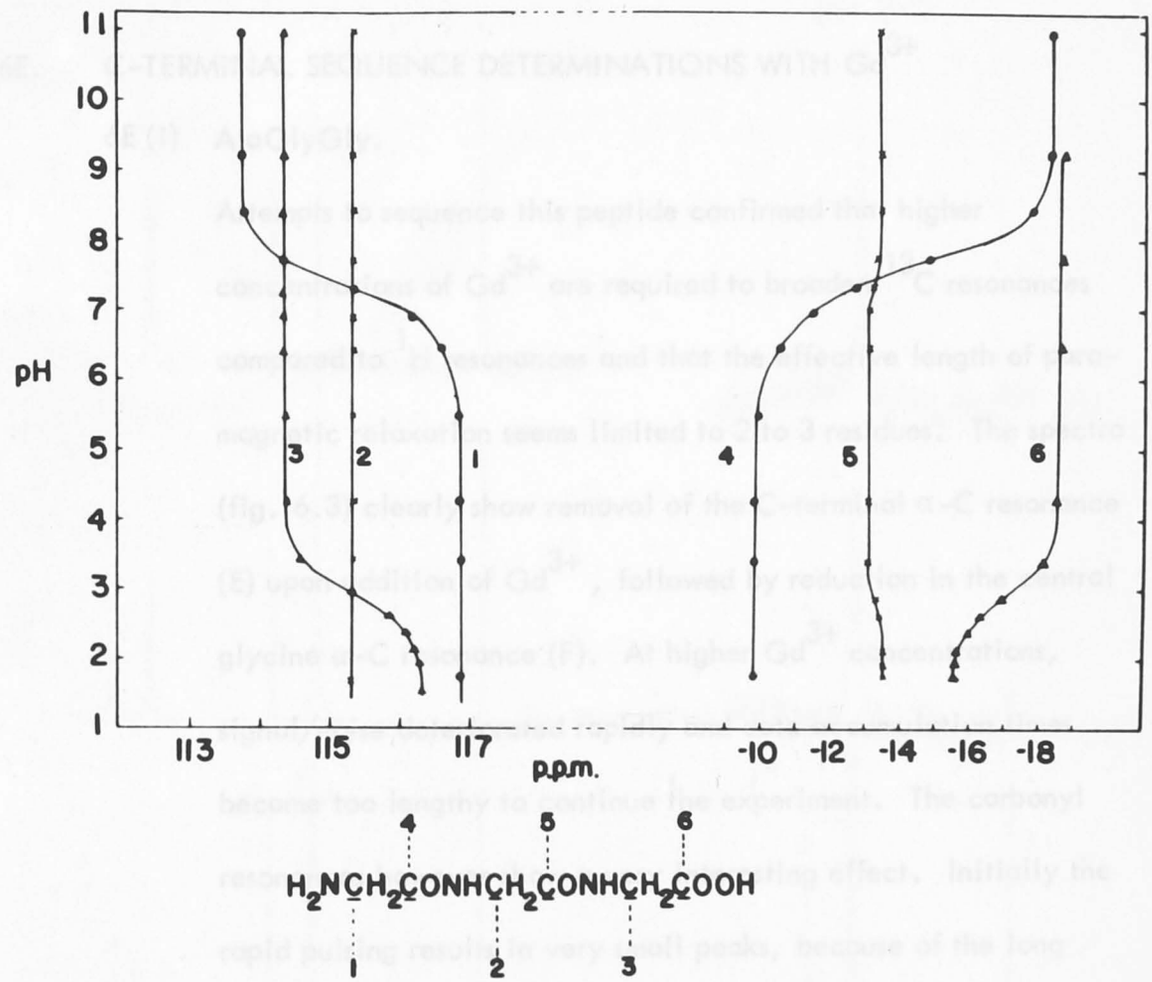


Fig. 6.2 pH shifts of ^{13}C NMR resonances of triglycine in 6M GuHCl shifts are recorded against an internal standard (GuHCl) in ppm. The numbers on each curve refer to the C nuclei denoted in the structural formula.

6D. EXAMPLES OF SEQUENCE DETERMINATION USING CMR

These examples are limited to T_2 (broadening) experiments with both Cu^{2+} and Gd^{3+}. Attempts to measure T_1 values proved impractical because of the very lengthy data accumulation times required.

6E. C-TERMINAL SEQUENCE DETERMINATIONS WITH Gd^{3+}

6E (i) AlaGlyGly.

Attempts to sequence this peptide confirmed that higher concentrations of Gd^{3+} are required to broaden ^{13}C resonances compared to ^{1}H resonances and that the effective length of paramagnetic relaxation seems limited to 2 to 3 residues. The spectra (fig. 6.3) clearly show removal of the C-terminal α-C resonance (E) upon addition of Gd^{3+}, followed by reduction in the central glycine α-C resonance (F). At higher Gd^{3+} concentrations, signal/noise deteriorated rapidly and data accumulation times became too lengthy to continue the experiment. The carbonyl resonances however show a very interesting effect. Initially the rapid pulsing results in very small peaks, because of the long relaxation times of these nuclei. As the Gd^{3+} ion concentration is increased broadening of the C-terminal glycine resonance (A) can clearly be seen but the central glycine resonance (C) and the alanine N-terminal resonance (B) show a sequential increase in peak height as their relaxation times undergo sequential reduction. This "reverse broadening" experiment is a useful confirmatory technique for sequence determination.

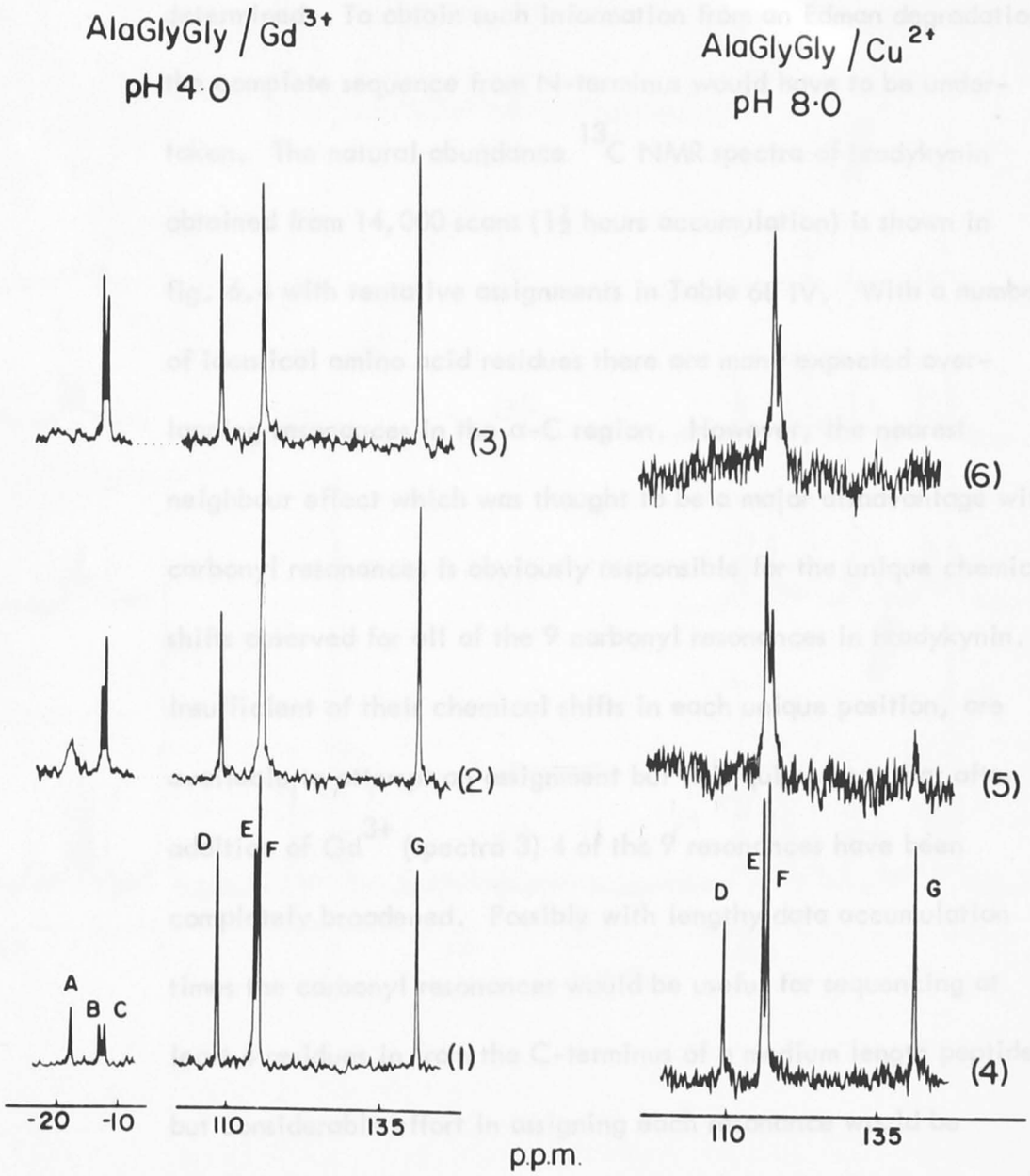


Fig. 6.3 ^{13}C spectra of AlaGlyGly at (a) pH 4.0 with Gd^{3+} addition and (b) pH 8.0 with Cu^{2+} addition. Assignments are A. C-terminal carboxyl, B. Alanine carbonyl, C. Central glycine carbonyl, D. Alanine α-C, E. C-terminal α-C, F. Central α-C, G. Alanine $\underline{C}H_3$.

6E(ii) Bradykynin (ArgProProGlyPheSerProPheArg).

The purpose of examining this peptide was to see if a sequence of 3 or 4 residues from the C-terminus of a long peptide could be determined. To obtain such information from an Edman degradation the complete sequence from N-terminus would have to be undertaken. The natural abundance ^{13}C NMR spectra of bradykynin obtained from 14,000 scans ($1\frac{1}{2}$ hours accumulation) is shown in fig. 6.4 with tentative assignments in Table 6E IV. With a number of identical amino acid residues there are many expected overlapping resonances in the α-C region. However, the nearest neighbour effect which was thought to be a major disadvantage with carbonyl resonances is obviously responsible for the unique chemical shifts observed for all of the 9 carbonyl resonances in bradykynin. Insufficient of their chemical shifts in each unique position, are available to attempt an assignment but it is quite clear that after addition of Gd^{3+} (spectra 3) 4 of the 9 resonances have been completely broadened. Possibly with lengthy data accumulation times the carbonyl resonances would be useful for sequencing at least 4 residues in from the C-terminus of a medium length peptide, but considerable effort in assigning each resonance would be required. Furthermore, the unique chemical shift of the carbonyl nuclei may be sufficient in itself to predict the neighbouring amino acid unit(s), particularly if high field spectrometers were used, but such an approach would require a very large number of reference

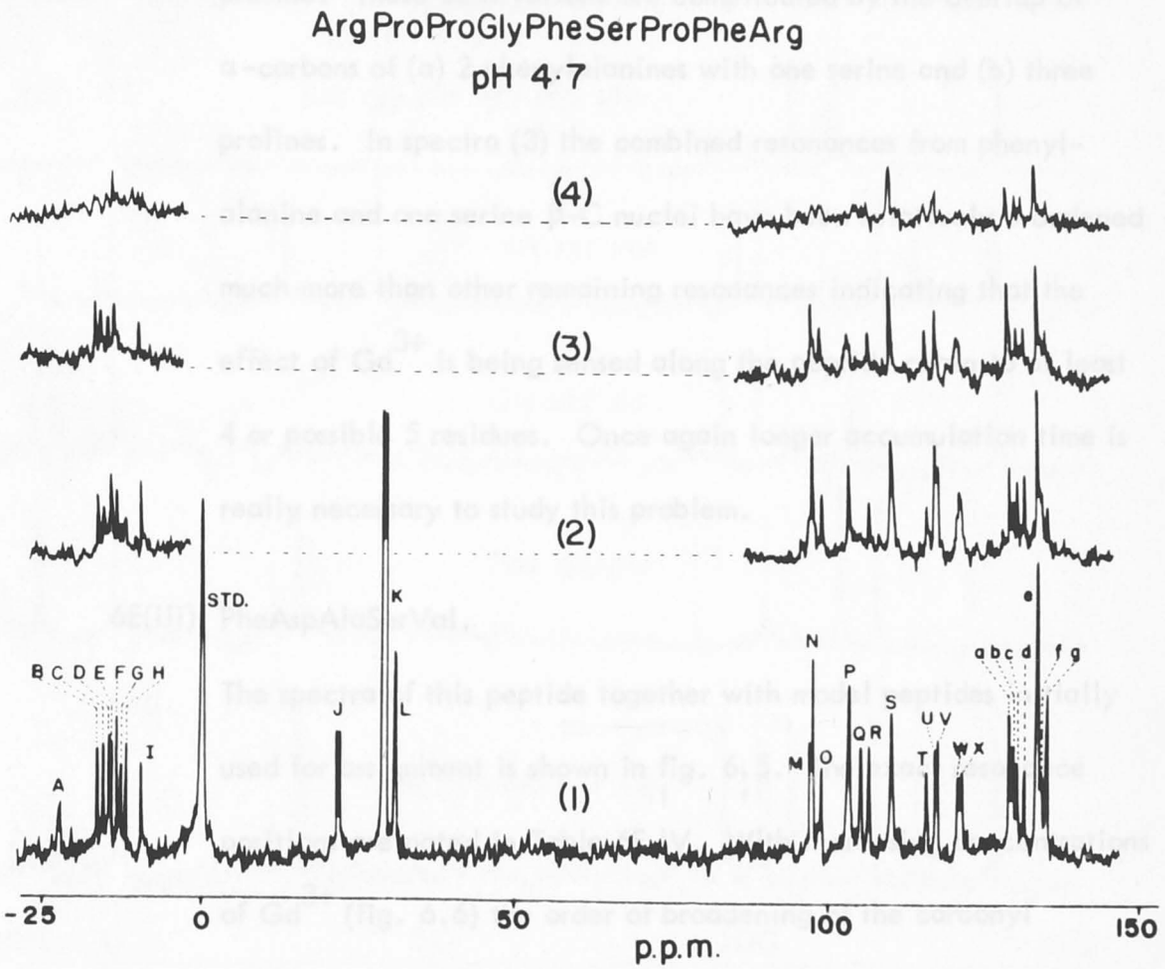


Fig. 6.4 ^{13}C spectra of Bradykynin at pH 4.0 in 6M GuHCl with additions of Gd^{3+} of (1) Nil, (2) 0.001M, (3) 0.005M, (5) 0.01 M. Assignments and chemical shifts are given in Table 6D IV.

peptides. Broadening of the α-C resonance is difficult to follow because of the overlapping resonances. After addition of Gd^{3+} the C-terminal α-C of arginine (5) disappeared completely, followed by selective drops in peak heights of the α-C of phenylalanine and proline. These observations are complicated by the overlap of α-carbons of (a) 2 phenylalanines with one serine and (b) three prolines. In spectra (3) the combined resonances from phenylalanine and one serine β-C nuclei have been relatively broadened much more than other remaining resonances indicating that the effect of Gd^{3+} is being sensed along the peptide chain to at least 4 or possible 5 residues. Once again longer accumulation time is really necessary to study this problem.

6E(iii) PheAspAlaSerVal.

The spectra of this peptide together with model peptides initially used for assignment is shown in fig. 6.5. The exact resonance positions are noted in Table 6E IV. With increasing concentrations of Gd^{3+} (fig. 6.6) the order of broadening of the carbonyl resonances occurs in sequence from the C-terminus along the complete peptide chain, but the order for α-C resonances is Val α-C + Ser β (which reduces greatly in height with first addition of Gd^{3+} but then more slowly) followed by Ser > Asp > Ala > Phe. This order is inverted between the third and fourth residues and clearly reflects some binding of Gd^{3+} to the side chain carboxyl group of the aspartyl residue. This is confirmed by the initial rapid broadening of the aspartyl βC resonance. Because the

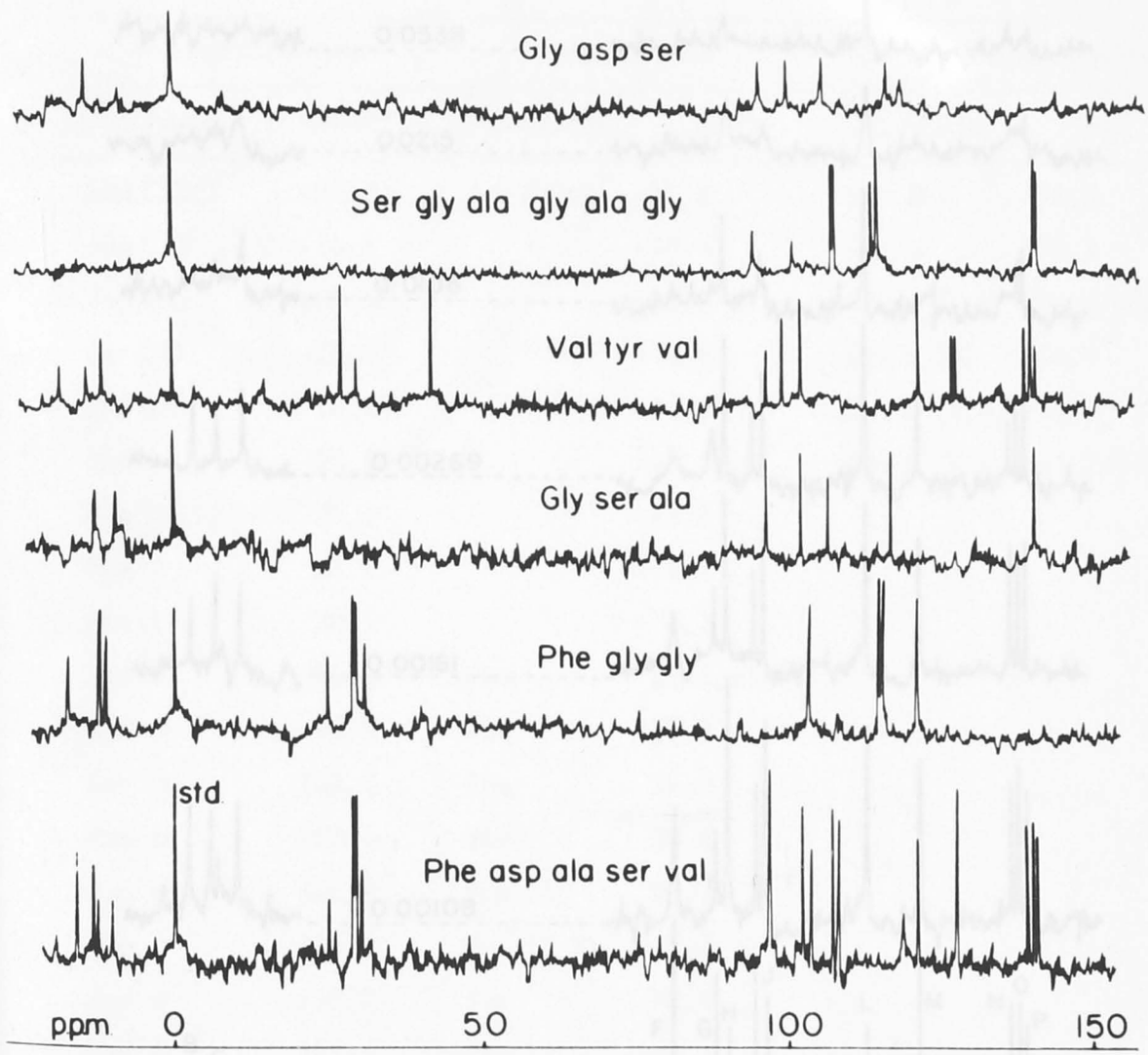


Fig. 6.5 ^{13}C NMR spectra of model peptides used to assign the resonances in the spectra of PheAspAlaSerVal. Peptides were dissolved in 6M GuHCl at pH 4.0 in water containing 5% D_2O. Chemical shifts are in ppm unpfield from the GuHCl resonance.

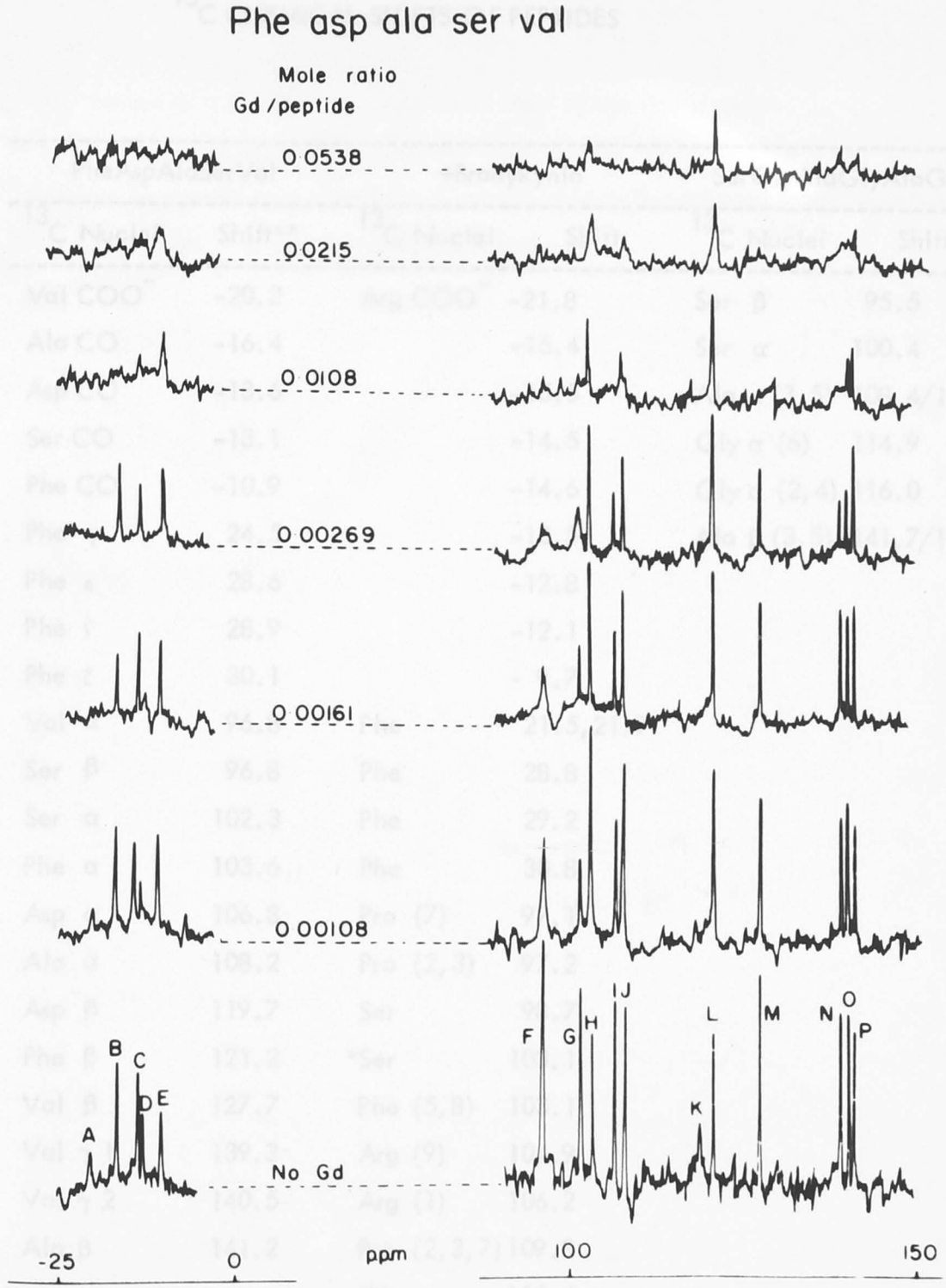


Fig. 6.6 CMR spectra of a 0.186 M solution of the pentapeptide PheAspAla-SerVal in the presence of increasing amounts of Gd^{3+} in 6M GuDCl at pH 4.0. The assignments of the resonances labeled A to P are based on the data in Table 6C I and are tabulated in Table 6E IV. Chemical shifts are in ppm upfield from GuHCl.

TABLE 6E IV

^{13}C CHEMICAL SHIFTS OF PEPTIDES

PheAspAlaSerVal		+Bradykynin		SerGlyAlaGlyAlaGly	
^{13}C Nuclei	Shift**	^{13}C Nuclei	Shift	^{13}C Nuclei	Shift
Val COO^-	-20.2	Arg COO^-	-21.8	Ser β	95.5
Ala CO	-16.4		-16.4	Ser α	100.4
Asp CO	-13.6		-15.5	Ala α (3,5)	108.4/108.9
Ser CO	-13.1		-14.5	Gly α (6)	114.9
Phe CO	-10.9		-14.6	Gly α (2,4)	116.0
Phe γ	24.5		-13.5	Ala β (3,5)	141.7/142.1
Phe ε	28.6		-12.8		
Phe ζ	28.9		-12.1		
Phe ζ	30.1		- 9.7		
Val α	96.8	Phe	21.5,21.9		
Ser β	96.8	Phe	28.8		
Ser α	102.3	Phe	29.2		
Phe α	103.6	Phe	30.8		
Asp α	106.8	Pro (7)	97.1		
Ala α	108.2	Pro (2,3)	97.2		
Asp β	119.7	Ser	98.7		
Phe β	121.2	*Ser	103.1		
Val β	127.7	Phe (5,8)	103.1		
Val γ 1	139.3	Arg (9)	104.9		
Val γ 2	140.5	Arg (1)	106.2		
Ala β	141.2	Pro (2,3,7)	109.8		
		Gly	115.4		
		Arg (1)	117.0		
		Arg (9)	117.1		
		Phe (5,8)	120.5,121.0		
		Pro (2,3,7)	128.4,129.1		
		Arg (9)	129.7		
		Arg (1)	130.8		

TABLE 6E IV

(continued)

* Noted as a shoulder on downfield edge of the Phe α peaks.

** Shifts are quoted in ppm upfield from GuHCl.

Numbers in parenthesis refer to each amino acid unit numbered from N-terminus.

\+ Sequence is ArgProProGlyPheSerProPheArg

correct broadening sequence was obtained for the carbonyl resonances but not for the α-C resonances can be explained by weaker side chain binding on the aspartyl residue than on the C-terminal valine residue. This weak side chain aspartyl binding is strong enough to broaden its α-C resonance (before that of the alanine α-C resonance) but not the more distant $\underline{C}$ = 0 resonance (before that of the alanine $\underline{C}$ = 0 resonance).

6F. N-TERMINAL SEQUENCE DETERMINATIONS WITH Cu^{2+}

In comparing the relative broadening effects of Cu^{2+} and Gd^{3+} (Section 6B), it was shown that an increasing concentration of Cu^{2+} could cause some selective broadening of the three α-C resonances of triglycine, although it was not possible to completely broaden these resonances. The following peptides were also examined to confirm that this result is general for other peptides and that the length of selective relaxation is of the order of three amino acid residues.

6F(i) AlaGlyGly.

A selection of many spectra recorded at increasing Cu^{2+} concentration (fig. 6.3) clearly shows that the relaxation effect can be sensed for at least three residues from the N-terminus, although the C-terminal glycine α-C resonance only shows slight broadening, with the highest copper concentration (spectra (6)). It was not possible to further increase the level of copper and still obtain spectra with acceptable noise levels using accumulation times of less than one hour.

6F (ii) SerGlyAlaGlyAlaGly.

^{13}C spectra and assignments are shown in fig. 6.7 and Table 6E IV respectively. The assignments for serine and glycine residues all follow logically from the chemical shifts previously obtained with model peptides (Table 6C I). In the case of the two central alanine resonances which show partially resolved peaks for both α-C and β-C nuclei, the tentative assignment given has been on the basis of this broadening experiment. As the Cu^{2+} concentration is increased both serine resonances (A & B) clearly broaden out completely followed by the adjacent glycine residue. The alanine α-C resonance (C) and β-C resonance (H) are the next peaks to show some broadening whilst at the same time there is a noticeable drop in the height of the remaining overlapping resonance of the two glycine α-carbons (F). The selective length of action of the Cu^{2+} probe is not clearcut in this case but appears to extend to at least the third and possibly the fourth residue from the N-terminus.

6G. ^{13}C ENRICHMENT OF NUCLEI IN PEPTIDES

The poor sensitivity of natural abundance ^{13}C NMR could obviously be overcome by enrichment of the ^{12}C nuclei with ^{13}C nuclei. Such a technique would have to introduce selective enrichment so that the simplicity of first order spectra be maintained without introducing the added complications of ^{13}C - ^{13}C spin spin splittings. In addition any reaction would have to be experimentally straightforward and be adaptable to all amino acid residues in peptides.

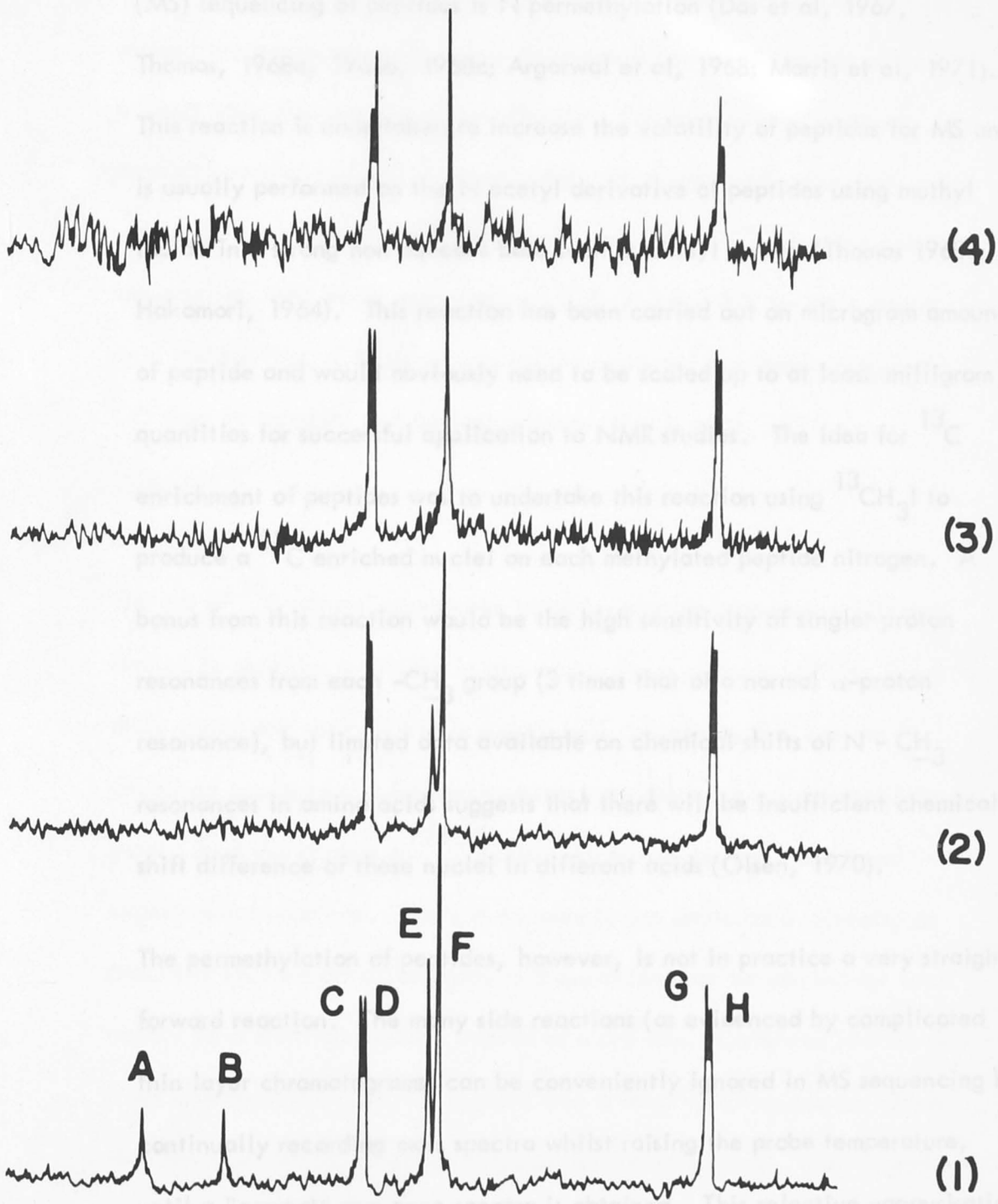


Fig. 6.7 Effect of Cu^{2+} on the ^{13}C NMR spectra of SerGlyAlaGlyAlaGly at Cu^{2+} concentrations of (1) Nil, (2) 0.005, (3) 0.02, (4) 0.09M. For peak assignments see test 6F(ii).

A chemical modification procedure which appears to fulfil these requirements and has already gained widespread use in mass spectrometry (MS) sequencing of peptides is N permethylation (Das et al, 1967, Thomas, 1968a, 1968b, 1968c; Argarwal et al, 1968; Morris et al, 1971). This reaction is undertaken to increase the volatility of peptides for MS and is usually performed on the N acetyl derivative of peptides using methyl iodide in a strong non aqueous base such as dimsyl sodium (Thomas 1968c; Hakomori, 1964). This reaction has been carried out on microgram amounts of peptide and would obviously need to be scaled up to at least milligram quantities for successful application to NMR studies. The idea for ^{13}C enrichment of peptides was to undertake this reaction using $^{13}CH_3I$ to produce a ^{13}C enriched nuclei on each methylated peptide nitrogen. A bonus from this reaction would be the high sensitivity of singlet proton resonances from each $-\underline{C}H_3$ group (3 times that of a normal α-proton resonance), but limited data available on chemical shifts of N - $\underline{C}H_3$ resonances in amino acids suggests that there will be insufficient chemical shift difference of these nuclei in different acids (Olsen, 1970).

The permethylation of peptides, however, is not in practice a very straightforward reaction. The many side reactions (as evidenced by complicated thin layer chromatograms) can be conveniently ignored in MS sequencing by continually recording mass spectra whilst raising the probe temperature, until a "correct" sequence spectra is obtained. This selective vapourisation of sample has been used for sequence determination of peptides of up to 10 residues (Morris et al, 1971) and seems to be a useful basis for sequencing peptide mixtures.

Attempts to scale up this reaction to modify 20 - 50 mg. of AcetylAla-AlaAlaOMe by the modified Hakomori procedure (Thomas 1968c) proved unsuccessful. Thin layer chromatograms of the resultant pale brown oil showed the presence of very many components and NMR spectra of the product were also extremely complicated. Presence of the "correct" product however was easily confirmed by the standard MS selective vapourisation method, which produced spectra at 70°C showing correct sequencing, with major mass peaks of 128, 213 and 329 for each amino acid residue (fig. 6.8) together with the expected mass peaks for loss of CO from N-terminal and central residues (Morris et al, 1971).

6H. CONCLUSION

The use of ^{13}C natural abundance NMR for peptide sequencing will be limited to peptides available in reasonably large amounts (100 mg.) and to those which are very soluble in 6M GuHCl. Assuming that one has unlimited access to a pulse spectrometer then there are many benefits to be gained from the change to ^{13}C nuclei. In particular there are fewer experimental problems. There is no need to use deuterated solvents, or NaOD or DCl for pH adjustment, and operations can be carried out in a normal laboratory atmosphere instead of a dry box. Compared to PMR there is no interference in the ^{13}C spectra from a large HDO resonance which requires much effort in repeated lyophilisation to remove. Most important is the change from multiplet α-proton resonances to singlet α-C resonances, which coupled with the additional chemical shift range, results in less overlap of ^{13}C resonances. Each peptide residue exhibits

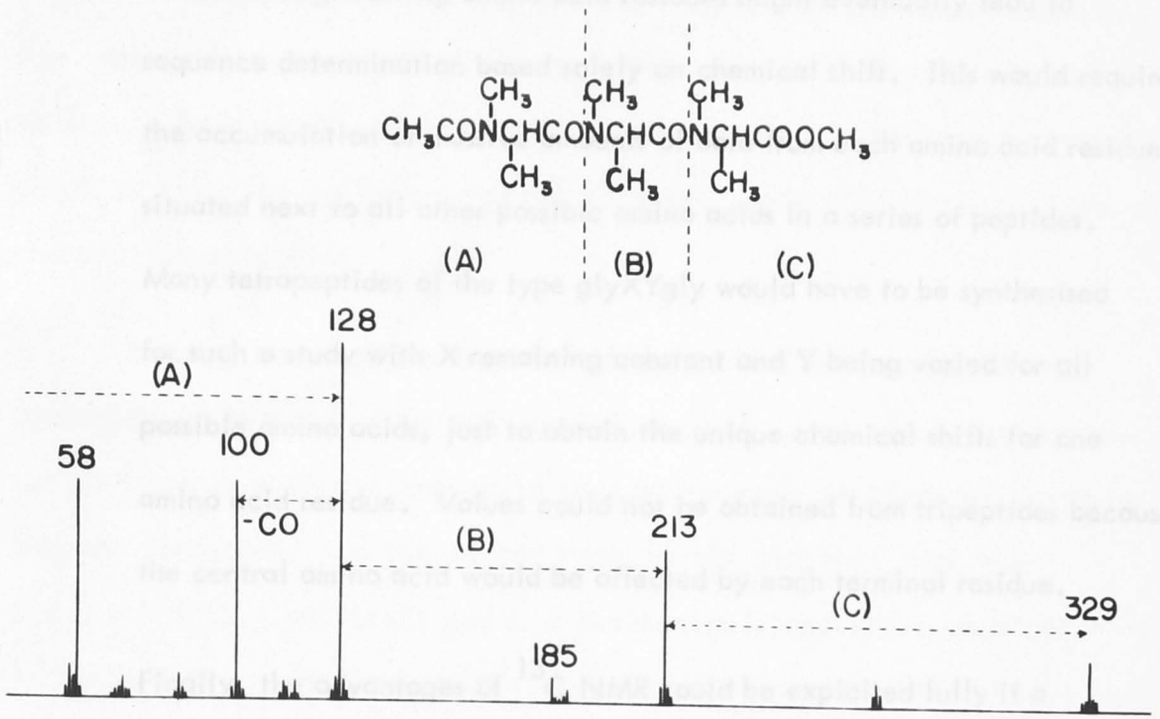


Fig. 6.8 Mass spectrum of N permethylated N acetyl trialanine methyl ester, obtained with a probe temperature of 70°C. Major mass peaks are labelled and fragments A, B and C refer to the three peptide fragments shown in the structural diagram.

two resonances from nuclei in the peptide backbone and assuming assignment difficulties for carbonyl resonances can be overcome, then both α-C and carbonyl resonances can be useful for sequence determination. The unique chemical shifts shown by carbonyl resonances caused by the effect of neighbouring amino acid residues might eventually lead to sequence determination based solely on chemical shift. This would require the accumulation of massive amounts of data from each amino acid residue situated next to all other possible amino acids in a series of peptides. Many tetrapeptides of the type glyXYgly would have to be synthesised for such a study with X remaining constant and Y being varied for all possible amino acids, just to obtain the unique chemical shifts for one amino acid residue. Values could not be obtained from tripeptides because the central amino acid would be affected by each terminal residue.

Finally, the advantages of ^{13}C NMR could be exploited fully if a successful enrichment procedure be developed together with improvements in binding sites and constants for Gd^{3+}.

CHAPTER 7

INCORPORATION OF NEW LANTHANIDE BINDING SITES INTO PEPTIDES

When Gd^{3+} ions are added to a peptide solution two effects are observed. Firstly Gd^{3+} ions bound to the carboxyl group cause a distance related relaxation of nuclei and secondly free Gd^{3+} ions cause a general relaxation of all nuclei. By increasing the ratio bound to free ions, the selective relaxation effect may be increased and hence it should be possible to sequence longer peptides. This may be brought about in practice by incorporating stronger lanthanide binding sites on the peptide molecule.

Binding processes between a substrate and lanthanide ion may conveniently be studied by NMR spectroscopy in cases where the substrate nuclei undergo either a change in relaxation rate or chemical shift on complex formation. In the latter case in its simplest form where $S + L \rightleftharpoons LS$ (S = substrate and L = lanthanide), and where chemical exchange between the bound and free substrate molecules is fast on the NMR time scale, only a single substrate resonance is observed, the chemical shift of which is the weighted average of the chemical shifts of substrate nuclei in the two environments. Assuming that the chemical shifts of the nuclei in the fully bound and free forms are known or can be independently measured, the position of the single averaged resonance will give a direct measure of the concentration of complex at equilibrium. This information may then be used for direct calculation of the binding constant.

There are three important properties of a complex that need to be ascertained:

(a) The stoichiometry of the complex(es).

(b) The chemical shift of the nuclei of the fully complexed substrate.

(c) The binding constant.

A knowledge of both (a) and (b) is essential to allow calculation of the binding constant by NMR spectroscopy.

7A. STOICHIOMETRY

Consider the formation of complexes LS and LS_2 formed from a substrate (S) and a lanthanide ion (L) according to the equilibria

$$L + S \overset{K_B}{\rightleftharpoons} LS \quad \text{and} \quad L + 2S \overset{K_{B2}}{\rightleftharpoons} LS_2$$

The one step binding constants K_B and K_{B2} for the complexes LS and LS_2 are given by:

$$K_B = \frac{[LS]}{[L]\,[S]} \ldots(1) \quad \text{and} \quad K_{B2} = \frac{[LS_2]}{[L]\,[S]^2} \ldots(2)$$

$[L]_o$ and $[S]_o$ will be used to denote total added concentrations of lanthanide and substrate respectively, whilst at equilibrium the concentrations of lanthanide, substrate and complexes will be given by [L], [S], [LS], and $[LS_2]$. Upon complexation the resonances of the substrate nuclei are shifted from their original position (H_o) by increments of ΔH up to a maximum value of ΔH_∞, representing the change in chemical shift due to complete complexation of all substrate molecules. Assuming that a particular nucleus from all substrate molecules exhibits the same chemical shift whether present in LS or LS_2 complexes, the averaged chemical shift of partially complexed substrate is given by:

$$\Delta H = \frac{[LS]\Delta H_\infty}{[S]_o} \quad \text{for 1:1 complexes} \tag{3}$$

$$\Delta H = \frac{([LS]+[LS_2])\Delta H_\infty}{[S]_o} \quad \text{for mixed complexes} \tag{4}$$

or $$\Delta H = \frac{[LS_2]\Delta H_\infty}{[S]_o} \quad \text{for solely 1:2 complexes} \tag{5}$$

Using this assumption and by restricting experiments to the conditions where $[S]_o >> [L]_o$, $[S]_o \leq \frac{1}{K_B}$ and $\Delta H << \Delta H_\infty$, Armitage et al (1972) have proposed a criterion for stoichiometry based on the equations:

$$[S]_o = [L]_o \Delta H_\infty (\frac{1}{\Delta H}) - (\frac{1}{K_B} + [L]_o) \quad \text{for LS complexes (derived in section 7Biii)} \tag{6}$$

$$[S]_o^2 - \frac{\Delta H_\infty}{\Delta H}[L]_o[S]_o + \frac{1}{K_B} = 0 \quad \text{for } LS_2 \text{ complexes} \tag{7}$$

They have demonstrated that a plot of $[S]_o$ vs $\frac{1}{\Delta H}$ at constant $[L]_o$ will be a straight line for 1:1 complexes as expected from equation (6). However it will exhibit marked curvature for 1:2 complexes (equation (7)), particularly at low substrate concentrations and/or small binding constants.

7A(i) Stoichiometry of the diglycine/Pr^{3+} system.

To study the stoichiometry of this system, which has been used to examine the binding capacity of unmodified peptides, a series of experiments were performed by varying diglycine concentration at fixed levels of $[Pr^{3+}]$, using the experimental constraints noted above.

Plots of $[S]_o$ vs $\frac{1}{\Delta H}$ at constant $[L]_o$ show straight lines (fig. 7.1a) which is good evidence for the formation of 1:1 complexes.

7A(ii) The chemical shift of the nuclei of the fully complexed substrate (ΔH_∞).

The basic relationship which has been used to calculate the concentration of a 1:1 complex in all NMR methods where fast exchange occurs is:

$$\Delta H = \frac{[LS]\Delta H_\infty}{[S]_o} \qquad (8)$$

Hence a knowledge of ΔH_∞, the change in chemical shift upon complete complexation of substrate molecules, is required to calculate [LS]. For the equilibrium

$$L + S \overset{K_B}{\rightleftharpoons} LS$$

$$K_B = \frac{[LS]}{[L][S]} = \frac{[LS]}{([L]_o-[LS])([S]_o-[LS])} \qquad (9)$$

Under conditions where $[S]_o >> [L]_o$ then $[S]_o-[LS] \doteq [S]_o$

By solving equation (9) for [LS], using this approximation and substituting into equation (8) for [LS] gives:

$$\Delta H = \frac{[L]_o \Delta H_\infty K_B}{1+K_B[S]_o} \qquad (10)$$

By using this equation it has become widespread practice to determine ΔH_∞ from the slope of a plot of ΔH vs $\frac{[L]_o}{[S]_o}$ (Rondeau et al, 1971; Demarco et al, 1970; Cockerill et al, 1970). However this method has a number of shortcomings. In the case where there is strong

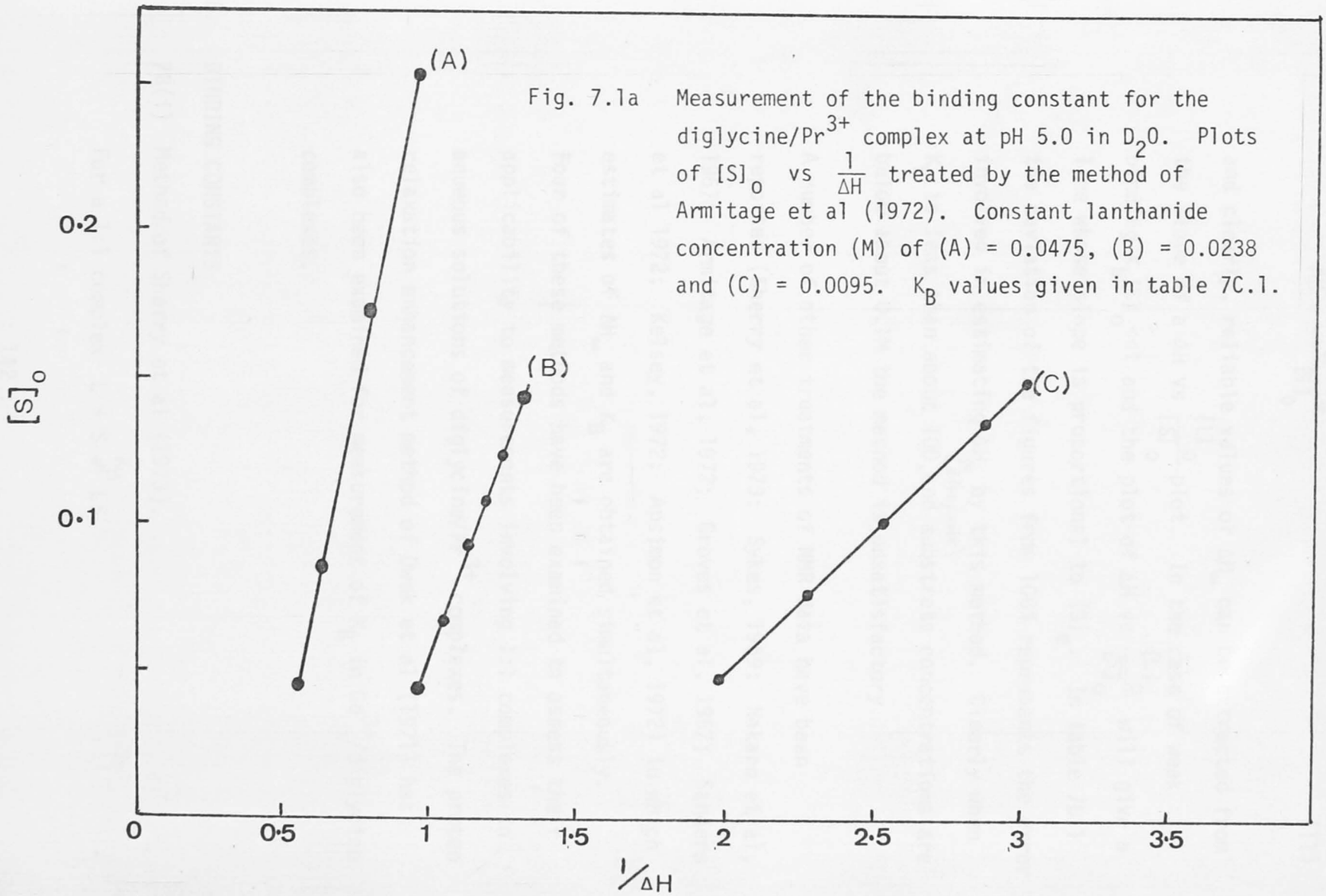


Fig. 7.1a Measurement of the binding constant for the diglycine/Pr^{3+} complex at pH 5.0 in D_2O. Plots of $[S]_o$ vs $\frac{1}{\Delta H}$ treated by the method of Armitage et al (1972). Constant lanthanide concentration (M) of (A) = 0.0475, (B) = 0.0238 and (C) = 0.0095. K_B values given in table 7C.1.

binding or high substrate concentrations $K_B[S]_o >> 1$ and equation (10) reduces to

$$\Delta H = \frac{\Delta H_\infty [L]_o}{[S]_o} \qquad (11)$$

and clearly, reliable values of ΔH_∞ can be extracted from the slope of a ΔH vs $\frac{[L]_o}{[S]_o}$ plot. In the case of weak binding $K_B[S]_o << 1$ and the plot of ΔH vs $\frac{[L]_o}{[S]_o}$ will give a line whose slope is proportional to $[S]_o$. In table 7B.1 the deviation of the figures from 100% represents the error involved in estimating ΔH_∞ by this method. Clearly when K_B is less than about 100 (litre/mole) and substrate concentrations are below about 0.1M the method is unsatisfactory.

A number of other treatments of NMR data have been reported (Sherry et al, 1973; Sykes, 1969; Nakano et al, 1967; Armitage et al, 1972; Groves et al, 1967; Sanders et al 1972; Kelsey, 1972; Apsimon et al, 1972) in which estimates of ΔH_∞ and K_B are obtained simultaneously. Four of these methods have been examined to assess their applicability to measurements involving 1:1 complexes in aqueous solutions of diglycine/Pr^{3+} complexes. The proton relaxation enhancement method of Dwek et al (1971) has also been examined for measurement of K_B in Gd^{3+}/diglycine complexes.

7B. BINDING CONSTANTS

7B(i) Method of Sherry et al (1973).

For a 1:1 complex $L + S \overset{K_B}{\rightleftharpoons} LS$

TABLE 7B 1

PERCENTAGE OF THEORETICAL SHIFT (ΔH_∞) OBTAINED FROM SLOPE OF ΔH vs $[L]_o/[S]_o$ PLOTS WITH ASSUMED SUBSTRATE CONCENTRATIONS AND K_B VALUES

Substrate Concentration (M)	% of theoretical ΔH_∞ calculated for K_B values of					
	1.0	10.0	20	50	100	1000 (litre mole^{-1})
0.371	27	79	88	95	97	100
0.186	16	65	79	90	95	100
0.037	4	27	43	65	79	97

$$K_B = \frac{[LS]}{[L][S]} \qquad [LS] = K_B[L][S] \tag{11a}$$

By substituting $[S]_o = [S] + [LS]$ in equation (8) gives

$$\Delta H = \frac{[LS]\Delta H_\infty}{[LS]+[S]} \tag{12}$$

and substituting for [LS] from (11a) in (12) gives

$$\Delta H = \left(\frac{K_B[L]}{1+K_B[L]}\right)\Delta H_\infty \quad \underline{or} \quad \frac{1}{\Delta H} = \frac{1}{K_B[L]\Delta H_\infty} + \frac{1}{\Delta H_\infty}$$

A plot of $\frac{1}{\Delta H}$ vs $\frac{1}{[L]}$ will give a straight line of intercept $\frac{1}{\Delta H_\infty}$ and a slopt of $\frac{1}{K_B \Delta H_\infty}$.

A FORTRAN program has been written (Appendix I) which initially sets $[L] = [L]_o$ and performs a least squares plot of $\frac{1}{\Delta H}$ vs $\frac{1}{[L]}$ to arrive at first estimates of K_B and ΔH_∞. Using these values new concentrations of [L] are calculated and the whole process repeated until convergence of K_B occurs. This method has failed to give consistent results with the diglycine/Pr^{3+} system and in many cases fails to converge on a value of K_B. The problems are thought to be due to the similarity in values for [L] and $[L]_o$ when K_B is small, and the method is not applicable to data obtained at constant $[L]_o$ with varying $[S]_o$. Hence it cannot be recommended for this system.

7B(ii) Method of Nakano et al (1967).

A graphical iteration procedure is used to obtain values for the equilibrium concentration of LS and hence K_B can be calculated. Since from equation (9)

$$K_B = \frac{[LS]}{([L]_o - [LS])([S]_o - [LS])}$$

$$[LS] = K_B[L]_o[S]_o - K_B[LS]([L]_o + [S]_o - [LS]) \qquad (13)$$

Substituting for [LS] from equation (8) and rearrangement gives

$$\frac{[L]_o}{\Delta H} = \frac{1}{\Delta H_\infty}([L]_o + [S]_o - [LS]) + \frac{1}{\Delta H_\infty K_B} \qquad (14)$$

A plot of $\frac{[L]_o}{\Delta H}$ vs $([L]_o + [S]_o)$ will give a line of slope approximately equal to $\frac{1}{\Delta H_\infty}$. This initial value of ΔH_∞ can then be used to calculate [LS] from equation (8). By substituting the values of [LS] in equation (14) and plotting a second graph of $\frac{[L]_o}{\Delta H}$ vs $([L]_o + [S]_o - [LS])$ an improved value of ΔH_∞ is obtained. These steps are repeated until convergence occurs. K_B can then be calculated from the value of ΔH_∞ and the intercept. In practice this is a very lengthy procedure and a FORTRAN program has been written (Appendix I) to compute K_B by this method. Consistent results have been obtained for the diglycine/Pr^{3+} system with experiments at constant $[L]_o$ and varying $[S]_o$ (table 7C.1).

7B(iii) Method of Armitage et al (1972).

For a 1:1 complex

$$K_B = \frac{[LS]}{[L][S]} = \frac{[LS]}{([L]_o - [LS])([S]_o - [LS])} \qquad (15)$$

If equation (15) is solved for [LS], and $[LS]^2$ term neglected, by substituting into equation (8)

$$\Delta H = \frac{K_B[L]_o \Delta H_\infty}{1 + K_B[S]_o + K_B[L]_o} \qquad (16)$$

which on rearranging gives

$$[S]_o = [L]_o \Delta H_\infty \left(\frac{1}{\Delta H}\right) - \left(\frac{1}{K_B} + [L]_o\right) \qquad (17)$$

Hence a plot of $[S]_o$ vs $\frac{1}{\Delta H}$, in experiments where $[S]_o$ is varied at constant $[L]_o$ will yield a straight line of slope $[L]_o \Delta H_\infty$ and intercept of $-\left(\frac{1}{K_B} + [L]_o\right)$.

The approximations used by Armitage to derive these equations are that $\Delta H << \Delta H_\infty$ and $\frac{1}{K_B} \geq [S]_o$, and experimental parameters must be selected to meet these conditions. In the diglycine/Pr^{3+} system where K_B is about 4 then $[S]_o$ should not exceed 0.25M. $[L]_o$ concentrations were arranged so that $\frac{\Delta H}{\Delta H_\infty} \leq 0.1$. Under these constraints excellent straight lines are obtained (fig. 7.1a) resulting in consistent values for K_B and ΔH_∞ (table 7C.1). Where such experimental constraints are acceptable this method seems reliable for measurement of K_B values. In cases where higher substrate concentrations are necessary and/or for sufficient accuracy it is only possible to measure ΔH values close to ΔH_∞ (as would be the case when ΔH_∞ is very small) then it becomes necessary to include the $[LS]^2$ term in the derivation and solve the resulting quadratic equation for [LS] according to the following method of Sykes (1969).

7B(iv) Method of Sykes (1969).

From the equation for 1:1 binding

$$K_B = \frac{[LS]}{([S]_o - [LS])([L]_o - [LS])}$$

rearranging gives

$$[LS]^2 - ([S]_o + [L]_o + \frac{1}{K_B})[LS] + [L]_o[S]_o = 0 \qquad (18)$$

solving for [LS] gives

$$[LS] = \frac{([S]_o + [L]_o + \frac{1}{K_B}) \pm \sqrt{([S]_o + [L]_o + \frac{1}{K_B})^2 - 4[L]_o[S]_o}}{2} \qquad (19)$$

The solution obtained for [LS] must be within the limits

$$0 \leq \frac{[LS]}{[L]_o} \leq 1$$

From equation (8) a plot of ΔH vs $\frac{[LS]}{[S]_o}$ should be a straight line of slope ΔH_∞. A FORTRAN program (Appendix I) has been written which computes [LS] over an incremented range of K_B values until the best fit for the ΔH vs $\frac{[LS]}{[S]_o}$ is obtained. For the diglycine/Pr^{3+} system this method gives very similar results to the simplified treatment of Armitage et al (1972) and in cases where the required experimental constraints are possible to use, the latter method, the full computer treatment, seems unnecessary. It should be noted that the computer based method of Sykes is not expected to be valid when K_B is very large,

because it simply becomes numerically insensitive when $\frac{1}{K_B}$ becomes much larger than $[S]_o$.

7B(v) Measurements of binding constants for Gd^{3+} using proton relaxation enhancement (PRE).

Approximate binding constants for the Gd^{3+}/Diglycine system may be inferred from the above studies with Pr^{3+}. It has not been possible to use the resonance broadening caused by Gd^{3+} ions for measurement of binding constants because of the difficulties in separating the amounts of broadening caused by free and bound Gd^{3+} ions. A direct measurement of K_B may be made however using the PRE method according to Dwek et al (1971). When binding occurs according to the equilibrium

$$L + S \overset{K_B}{\rightleftharpoons} LS$$

water protons in the system may exist in either of three environments, (a) in the bulk aqueous solvent; (b) bound to free Gd^{3+} ions; or (c) bound to Gd^{3+} in the complex. Since the solvent is present in a large molar excess then exchange of protons between free Gd^{3+} ions and the complex may be neglected and only exchange between these centres and the bulk solvent need be considered. The paramagnetic contribution to the observed spin lattice relaxation rate $(\frac{1}{T_{1P}^*})$ for the water protons is given by

$$\frac{1}{T_{1P}^*} = (\frac{1}{T_{1P}^*})_f + (\frac{1}{T_{1P}^*})_b \qquad (20)$$

Both T_1 and T_2 values are of course affected by the paramagnetic ion, but for simplicity only T_1 values are considered, because this is the quantity required in PRE experiments. In equation (20) f and b denote water protons bound to free Gd^{3+} and those bound to complexed Gd^{3+} respectively. An * denotes the presence of a macromolecule. Also,

$$\frac{1}{T^*_{1P}} = \frac{1}{T^*_{1,obs}} - \frac{1}{T^*_{1,0}} \qquad (21)$$

where $\frac{1}{T^*_{1,obs}}$ is the observed spin lattice relaxation rate in presence of Gd^{3+} and $\frac{1}{T^*_{1,0}}$ the rate in the absence of Gd^{3+}.

From equation (21) and equations for T_1 on pages 14 and 18

$$\frac{1}{T^*_{1P}} = \left(\frac{P_M q}{T^*_{1M} + \tau^*_M}\right)_f + \left(\frac{P^*_M q^*}{T^*_{1M} + \tau^*_M}\right)_b \qquad (22)$$

where q and q* are the number of water molecules in the first hydration sphere of the metal ion when in the free and bound forms. P_M and P_M* are the mole fractions of paramagnetic ion in free and bound states.

The normal empirical quantity used to define this relaxation effect is termed the relaxation enhancement factor (ε*) where

$$\varepsilon^* = \frac{\frac{1}{T^*_{1P}}}{\frac{1}{T_{1P}}} = \frac{\frac{1}{T^*_{1,obs}} - \frac{1}{T^*_{1,0}}}{\frac{1}{T_{1,obs}} - \frac{1}{T_{1,0}}} \qquad (23)$$

By neglecting outer sphere contributions to the relaxation rate and combining equation (21) and (22) gives

$$\frac{1}{T^*_{1,obs}} - \frac{1}{T^*_{1,0}} = \frac{M_{fq}}{N_p}\left(\frac{1}{T^*_{1M} + \tau^*_M}\right)_f + \frac{M_{bq^*}}{N_p}\left(\frac{1}{T^*_{1M} + \tau^*_M}\right)_b \qquad (24)$$

where M_b and M_f are the molar concentrations of metal bound and free. N_p is the molar concentration of protons in solution. A similar equation to (24) exists when no macromolecule is present, i.e. identical equation without * values.

Substituting these equations into equation (23) gives

$$\varepsilon^* = \frac{M_b}{M_t}\varepsilon_b + \frac{M_f}{M_t}\varepsilon_f \qquad (25)$$

where M_t is the total metal ion concentration and

$$\varepsilon_f = \frac{(T_{1M} + \tau_M)_f}{(T^*_{1M} + \tau^*_M)_f} \quad \text{and} \quad \varepsilon_b = \frac{q^*}{q}\frac{(T_{1M} + \tau_M)_f}{(T^*_{1M} + \tau^*_M)_b}$$

ε_b has been termed the characteristic enhancement of the metal macromolecule complex. Equation (25) may conveniently be written as

$$\varepsilon^* = x_b\varepsilon_b + x_f\varepsilon_f \qquad (26)$$

where x_b and x_f are the mole fractions of bound and free metal ions.

To apply this phenomenon to binding studies, consider the equilibrium $L + S \overset{K_B}{\rightleftharpoons} LS$

$$\frac{1}{K_B} = \frac{[S]_f [L]_f}{[LS]} = \frac{[S]_f([L]_t - [L]_b)}{[L]_b} \qquad (26a)$$

where subscripts t, f and b refer to total, free and bound metal or substrate. From equation (26)

$$\varepsilon^* = x_b \varepsilon_b + (1 - x_b) \qquad (27)$$

this assumes that $\varepsilon_f = 1$, because the microviscosity of the regions surrounding the free paramagnetic ions is identical in solutions with or without macromolecules (Dwek, 1973).

From equation (26a)

$$x_b = \frac{[L]_b}{[L]_t} = \frac{[S]_f}{\frac{1}{K_B} + [S]_f} \qquad (28)$$

And from equation (27)

$$x_b = \frac{\varepsilon^* - 1}{\varepsilon_b - 1}$$

therefore

$$\frac{1}{\varepsilon^* - 1} = \frac{\frac{1}{K_B} + [S]_f}{[S]_f(\varepsilon_b - 1)} \qquad (29)$$

$$\frac{1}{\varepsilon^* - 1} = \frac{1}{[S]_f}\left(\frac{\frac{1}{K_B}}{(\varepsilon_b - 1)}\right) + \left(\frac{1}{\varepsilon_b - 1}\right) \qquad (30)$$

At high concentrations, $[S]_t \approx [S]_f$ and therefore a plot of $\frac{1}{\varepsilon^* - 1}$ vs $\frac{1}{[S]_t}$ will yield a straight line of slope

$\frac{1}{K_B} / (\varepsilon_b - 1)$ and intercept $1/(\varepsilon_b - 1)$

TABLE 7C 1

COMPARISON OF K_B VALUES OBTAINED FOR

BINDING OF Pr^{3+} TO DIGLYCINE AT pH 5.0 IN D_2O SOLUTION

Method	Range of Substrate Concentration (M)	Lanthanide Concentration (M)	Binding Constant K_B (ℓ mole^{-1})	Maximum Shift(ppm) ΔH_∞
**Armitage (1972)	0.3419 - 0.0427	0.0475	4.3	11.8
	0.1416 - 0.0425	0.0238	4.2	11.5
	0.1641 - 0.0492	0.0095	4.4	11.7
Sykes (1969)	0.3419 - 0.0427	0.0475	4.5	11.9
	0.1416 - 0.0425	0.0238	4.3	11.6
	0.1641 - 0.0492	0.0095	4.5	11.6
Nakano (1967)	0.3419 - 0.0427	0.0475	4.3	11.8
	0.1416 - 0.0425	0.0238	5.4	10.9
	0.1641 - 0.0492	0.0095	5.6	10.8
Dwek (1973)*	0.050 - 0.400	0.001	1.6	-

Literature values for similar compounds. Alanine/Nd^{3+}, K_B = 4.4 by titration; Histidine/Nd^{3+} K_B = 1.8; Serine/Nd^{3+}, K_B = 9.7 (Sherry et al 1973). Glycine/Eu^{3+}, K_B = 4.1; Acetate/Pr^{3+}, K_B = 66 (Sillen, L.G., Equilibrium Constants, Chemical Society special publication, part I, 1971).

* PRE method for Gd^{3+}/diglycine at pH 5.0 in D_2O.

** Values obtained from a least squares plot (fig. 7.1a).

An iterative computer program has been written (Appendix I) which increments values of K_B until the best straight line fit to this plot is obtained.

7C. DISCUSSION

There are two main practical approaches to obtain shift data suitable for measurement of equilibrium constants. The substrate may be held constant with varying lanthanide or the lanthanide held constant with varying substrate. Data obtained under the former conditions gave widely varying results by all of the methods examined and in many cases no convergence was obtained at all. As noted by Armitage et al (1972) wide variations in lanthanide concentrations cause large changes in magnetic susceptibility which can result in general changes in chemical shifts which cannot be adequately corrected with an internal standard. With data obtained from experiments using constant lanthanide concentration with varying substrate consistent results were obtained (Table 7C.1), although the method of Nakano did tend to give higher K_B values and lower ΔH_∞ values at very low lanthanide concentrations. This is probably due to lack of numerical sensitivity when values of $[S]_o + [L]_o$ are similar to $[S]_o + [L]_o - [LS]$.

For the diglycine/Pr^{3+} system there seems little need to resort to the full computer treatment of data, whilst the graphical method of Armitage et al (1972) gives comparable results from a simple $[S]_o$ vs $\frac{1}{\Delta H}$ plot. A computed

solution to the full quadratic equation according to Sykes (1969) would be necessary when it is experimentally impossible to obtain accurate data under the Armitage constraints. For example when ΔH_{∞} is very small, it would not be possible to measure accurate values of ΔH, since ΔH would be immeasurably small under the condition that $\Delta H << H_{\infty}$. The values of K_B for the diglycine/Pr^{3+} complex obtained from the Armitage and Sykes methods are between 4.2 and 4.5 ℓ mole^{-1}. No literature values could be obtained for this complex; similar values have been measured by potentiometric method for alanine/Nd^{3+} (K_B = 4.4) and glycine/Eu^{3+} (K_B = 4.1) (see footnote table 7C.1).

The value of K_B = 1.6 for the Diglycine/Gd^{3+} system is lower than the value obtained for the Pr^{3+} complex. The difference could be due to errors arising from the PRE method, particularly in the measurement of T_1 values, but is more likely due to the change in lanthanide ion. In comparing the binding constants of acetate to each of the lanthanide ions, Kolat et al (1962) demonstrated considerable variation across the lanthanide series. For example K_B values for acetate with Pr^{3+} and Gd^{3+} were 160 and 140 respectively (which seem surprisingly high for a single carboxyl group). Similar trends have also been noted with oxalate (Graffeo et al, 1968) and anthranilate (Silber et al, 1969) and it therefore seems probable

that the difference in values between Gd^{3+} and Pr^{3+} binding to diglycine is acceptable.

7D. CHEMICAL MODIFICATIONS

7D(i) Dicarboxymethylation.

Carboxymethylation of amino acids, peptides and proteins with iodo acetate is a well established procedure (Stark, 1970) and fairly detailed methods have been published for amino acids (Korman and Clark, 1956). Binding constants of Gd^{3+} to various dicarboxylic acids are usually at least an order of magnitude greater than for monocarboxylic acids. Therefore attachment of a dicarboxymethyl

group to the N-terminus of a peptide should provide a unique binding site for Gd^{3+} or Pr^{3+} which may well allow a sequence determination from the N-terminus. Disubstitution on the terminal amino group can be achieved by using a large excess of iodoacetate at pH 9.0 (Korman and Clarke, 1956). A peptide modified in this way would then posses carboxyl groups on both ends of the peptide chain, but little competition for Gd^{3+} binding at the C-terminal carboxyl group should result if the dicarboxymethyl group has a binding constant orders of magnitude greater.

7D(i) A. Experimental.

The conditions of carboxymethylation recommended by Korman and Clarke (1956) of excess bromo acetate at pH 9 for 3 days at 36°C, were found to be grossly in excess of those actually needed for complete reaction (iodo acetate pH 9, 1 hour at 35°C). In all cases a 5 fold molar excess of iodo acetate over peptide was used and the course of reaction was followed by the uptake of NaOH in a pH-stat maintained at 9.0. The lengthy isolation procedure of Korman and Clarke (1956) using mercuric salts was also dispensed with because the nitrogen analyses of their product were not particularly close to theory. In fact the sole criteria for confirmation of product was by identification of a modified N-terminal amino acid after complete

hydrolysis of the reacted peptide. A more convenient procedure was found to be as follows. The reaction mixture was acidified to pH 2 and the complete sample lyophilised under vacuum. The crude product and excess iodo acetate were dissolved in a minimum quantity of water and the solution poured slowly into a large volume of absolute alcohol with vigorous stirring. Precipitation of the required derivative in a fine granular form resulted, whilst excess iodo acetate remained in solution. This procedure was successfully used to prepare derivatives of diglycine, triglycine, tetraglycine, and triglycine amide. Attempts to prepare diCMB triglycine methyl ester resulted in hydrolysis of the ester group. Purity of the diCBM derivatives was confirmed by thin layer chromatography using standard silica gel G layers with a solvent system of n-butanol/water/acetic acid (5 : 4 : 1), and by their NMR spectra and associated pH resonance shifts (fig. 7.2).

B. Binding studies with Gd^{3+} and Pr^{3+} on dicarboxymethylated peptides.

Attempts to use the preceding methods to measure Pr^{3+} / diCBM peptide binding constants were unsuccessful. Addition of Pr^{3+} to these derivatives caused general resonance broadening and no shifting. Even more surprising was the fact that this non specific broadening occurred not

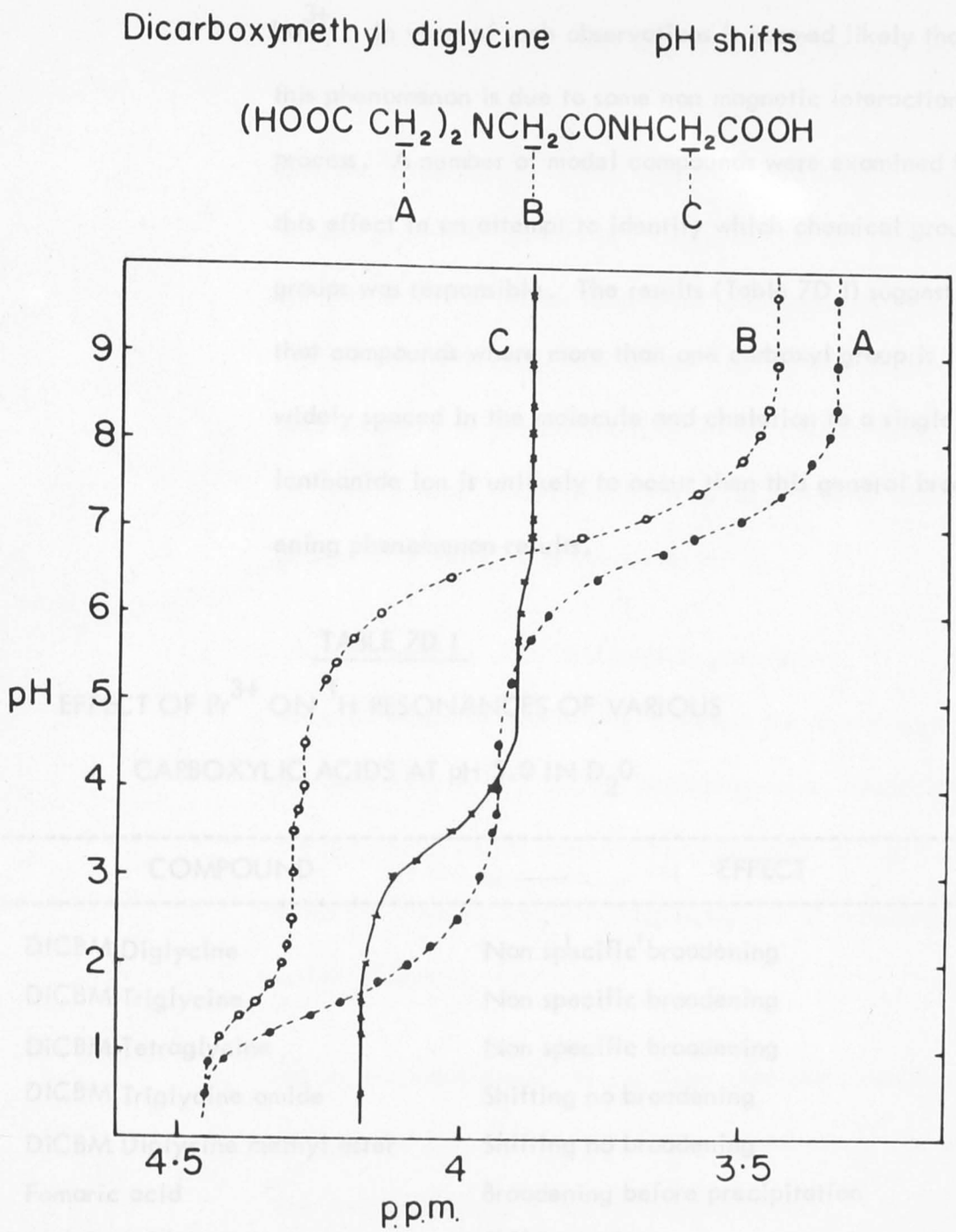


Fig. 7.2 pH shifts of ^{1}H resonances of dicarboxymethyldiglycine in D_2O, against an external standard of TMS.

only with Pr^{3+} and Gd^{3+} but even with the diamagnetic La^{3+}. In view of such observations it seemed likely that this phenomenon is due to some non magnetic interaction or process. A number of model compounds were examined for this effect in an attempt to identify which chemical group or groups was responsible. The results (Table 7D I) suggest that compounds where more than one carboxyl group is widely spaced in the molecule and chelation to a single lanthanide ion is unlikely to occur then this general broadening phenomenon results.

TABLE 7D I

EFFECT OF Pr^{3+} ON ^1H RESONANCES OF VARIOUS CARBOXYLIC ACIDS AT pH 5.0 IN D_2O

COMPOUND	EFFECT
DICBM Diglycine	Non specific broadening
DiCBM Triglycine	Non specific broadening
DiCBM Tetraglycine	Non specific broadening
DICBM Triglycine amide	Shifting no broadening
DiCBM Diglycine methyl ester	Shifting no broadening
Fumaric acid	Broadening before precipitation
Maleic acid	Shifting only
E.D.T.A.	Broadening no shifting
Adipic acid	Some shifting and broadening followed by precipitation.

The comparison between fumaric and maleic acids (fig. 7.3) is an example of chemically similar compounds one of which sterically favours chelation, and one which cannot chelate to a single Pr^{3+} ion. Fumaric acid is more likely to form polymeric species in solution as it is not possible for both carboxyl groups to chelate to a single Pr^{3+} ion. Conceivably this could also be the case for the DiCBM peptide derivatives containing an unmodified C-terminal carboxyl group. (Note that with amide and ester blocking groups no broadening is caused by Pr^{3+}).

C. Viscosity studies of DiCBM peptide derivatives.

Solution viscosity studies should enable the detection of any polymeric Pr^{3+} / DiCBM peptide species, except where the polymers exist as non-solvated rigid spheres. In such cases the intrinsic viscosity is independant of molecular weight (Einstein, 1906).

Among the many empirical equations proposed to explain the variation of specific viscosity with concentration, the one most commonly used is due to Huggins (1942)

$$\eta sp/C = [\eta] + K^1 [\eta]^2 C$$

where K^1 is the Huggins constant and has been related to the shapes of molecules in solution. Plots of $\eta sp/C$ vs C allow the value of $[\eta]$ to be obtained from the graph extrapolated to zero concentration.

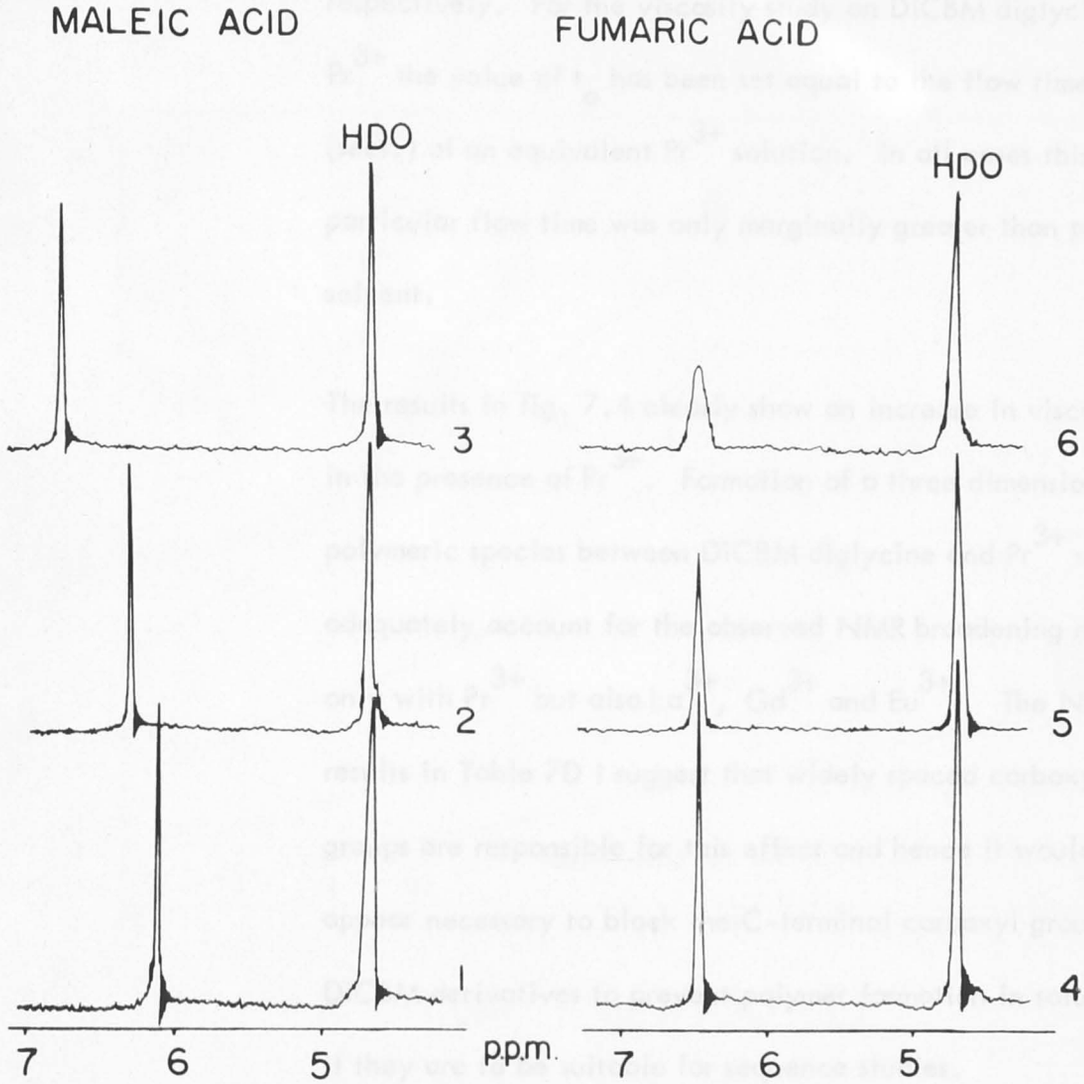


Fig. 7.3 Continuous wave 100 MHz PMR spectra of 0.1M solutions of Maleic and Fumaric acids in D_2O at pH 5.0 before and after addition of Pr^{3+} concentrations of 1. Nil, 2. 0.01M, 3. 0.03M, 4. Nil, 5. 0.009M, 6. 0.02M. Chemical shifts are quoted in ppm downfield from external TMS.

$$\eta_{sp} = (t - t_o) / t_o$$

where t and t_o are flow times (secs.) of solution and solvent respectively. For the viscosity study on DiCBM diglycine/ Pr^{3+} the value of t_o has been set equal to the flow time (secs.) of an equivalent Pr^{3+} solution. In all cases this particular flow time was only marginally greater than pure solvent.

The results in fig. 7.4 clearly show an increase in viscosity in the presence of Pr^{3+}. Formation of a three dimensional polymeric species between DiCBM diglycine and Pr^{3+} would adequately account for the observed NMR broadening not only with Pr^{3+} but also La^{3+}, Gd^{3+} and Eu^{3+}. The NMR results in Table 7D I suggest that widely spaced carboxyl groups are responsible for this effect and hence it would appear necessary to block the C-terminal carboxyl group of DiCBM derivatives to prevent polymer formation in solution, if they are to be suitable for sequence studies.

D. Suitable carboxyl blocking groups for peptide derivatives. The need to block the C-terminal carboxyl (and possibly any other side chain carboxyl groups) prior to attachment of a binding site to the N-terminal is obvious from the preceding viscosity studies. Esters and amides of peptides have been examined using T_1 measurement to determine the

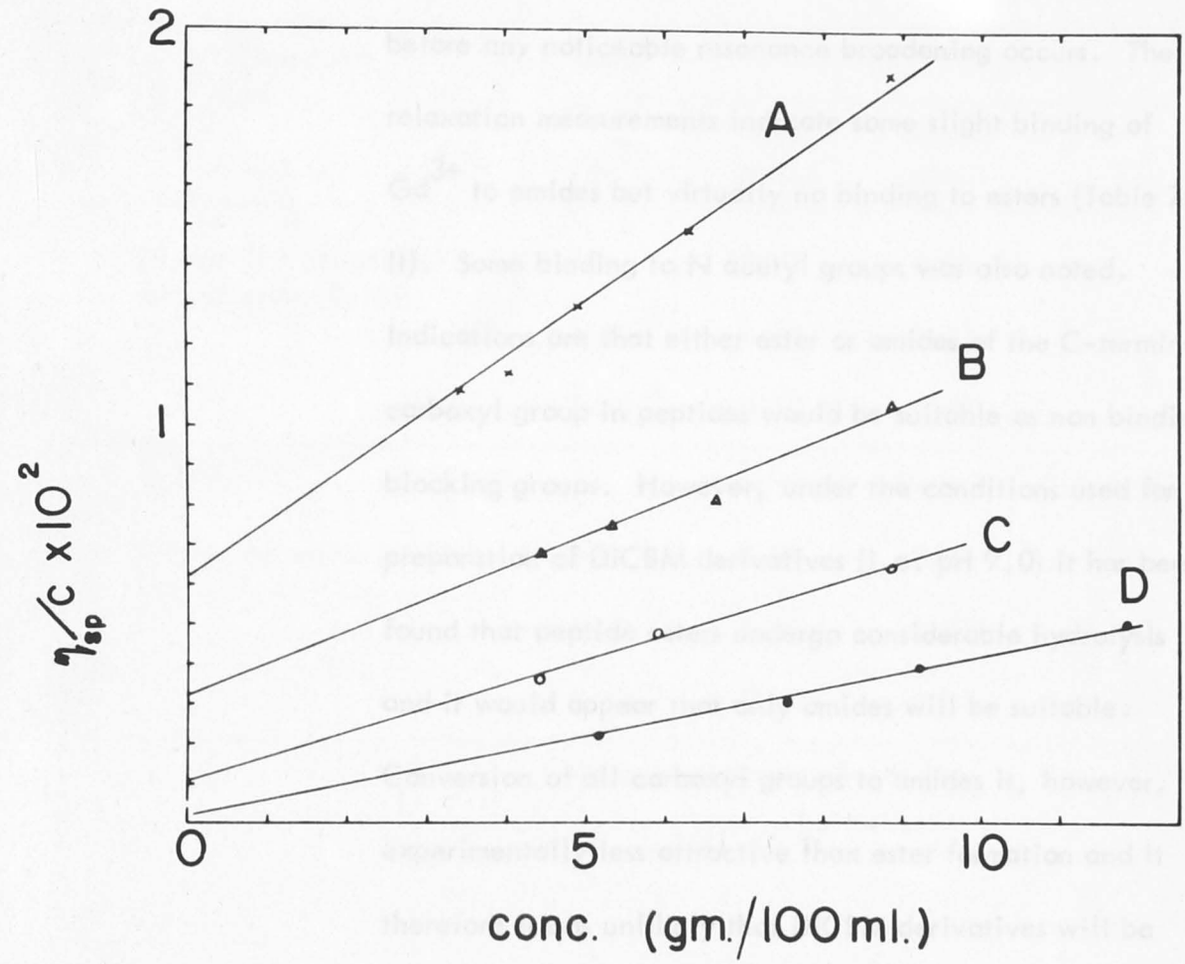


Fig. 7.4 Plots of viscosity data according to the Huggins equation (1942) for intrinisic viscosity measurements with Pr^{3+} / dicarboxymethyl-diglycine ratios of A. 3.0 B. 1.5 C. 0.8 D. 0.4 which result in $[\eta]$ values of A. 6.3×10^{-2}, B. 3.3×10^{-2}, C. 1.1×10^{-2} and D. 0.2×10^{-2}.

extent of Gd^{3+} binding on either of these moeities. T_1 measurements are a very sensitive probe to detect binding of paramagnetic ions on adjacent groups, particularly with Gd^{3+}, where spin lattice relaxation times are reduced even before any noticeable resonance broadening occurs. The relaxation measurements indicate some slight binding of Gd^{3+} to amides but virtually no binding to esters (Table 7D II). Some binding to N acetyl groups was also noted. Indications are that either ester or amides of the C-terminal carboxyl group in peptides would be suitable as non binding blocking groups. However, under the conditions used for preparation of DiCBM derivatives (i.e. pH 9.0) it has been found that peptide esters undergo considerable hydrolysis and it would appear that only amides will be suitable. Conversion of all carboxyl groups to amides is, however, experimentally less attractive than ester formation and it therefore seems unlikely that DiCBM derivatives will be useful for sequence determination.

TABLE 7D II

EFFECT OF Gd^{3+} ON THE ^1H NUCLEI T_1 VALUES

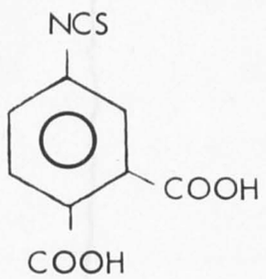
OF PEPTIDE DERIVATIVES

Compound	Molar Conc.^{n} Gd^{3+}	T_1 (secs.) of protons		
		Ester group	Acetyl group	
N Acetyl tri alanine - methyl ester (0.1M)	NIL	0.99	0.93	
	4×10^{-5}	0.96	0.81	
	2×10^{-4}	0.97	0.69	
		N-terminal	Central	C-terminal
Triglycine amide (0.1M)	NIL	0.90	0.93	0.94
	4×10^{-5}	0.92	1.02	1.06
	2×10^{-4}	0.90	0.94	1.05

7E. ISOTHIOCYANATE DERIVATIVES

Isothiocyanates react specifically with the terminal amino group of peptides (Stark, 1970). Reaction with certain isothiocyanates which themselves contain a lanthanide binding site, may provide a unique method for incorporating an N-terminal lanthanide binding site into peptides. In attempting such a modification the isothiocyanate of phthalic acid (I) has been prepared (Howell, 1976)

(I)



This compound reacts with triglycine at pH 8.0 in approximately 15 minutes at room temperature. The reaction may be followed by NMR as the N-terminal α-proton resonance shifts downfield by about 1 ppm during reaction. However the resulting product appears to form very strong complexes with Pr^{3+} and Gd^{3+} which readily precipitate from aqueous solution and prevent their use for binding or shifting studies in solution.

Reaction of (I) with triglycine amide produces the expected product, again with the N-terminal α-proton resonance showing a downfield shift. (fig. 7.5). Gd^{3+} exerts a sequential broadening effect from the N-terminus as expected, but the reduction in T_1 values is not sequential from the N-terminus (Table 7E I). This is most likely due to the formation of complexes other than the 1 : 1 at the low Gd^{3+} concentrations used for T_1 measurement. With complexes containing a single Gd^{3+} ion and 2, 3, 4

Fig. 7.5 1. PMR spectra of Triglycine amide.

2. PMR spectra of reaction product with isothiocyanate of phthalic acid.

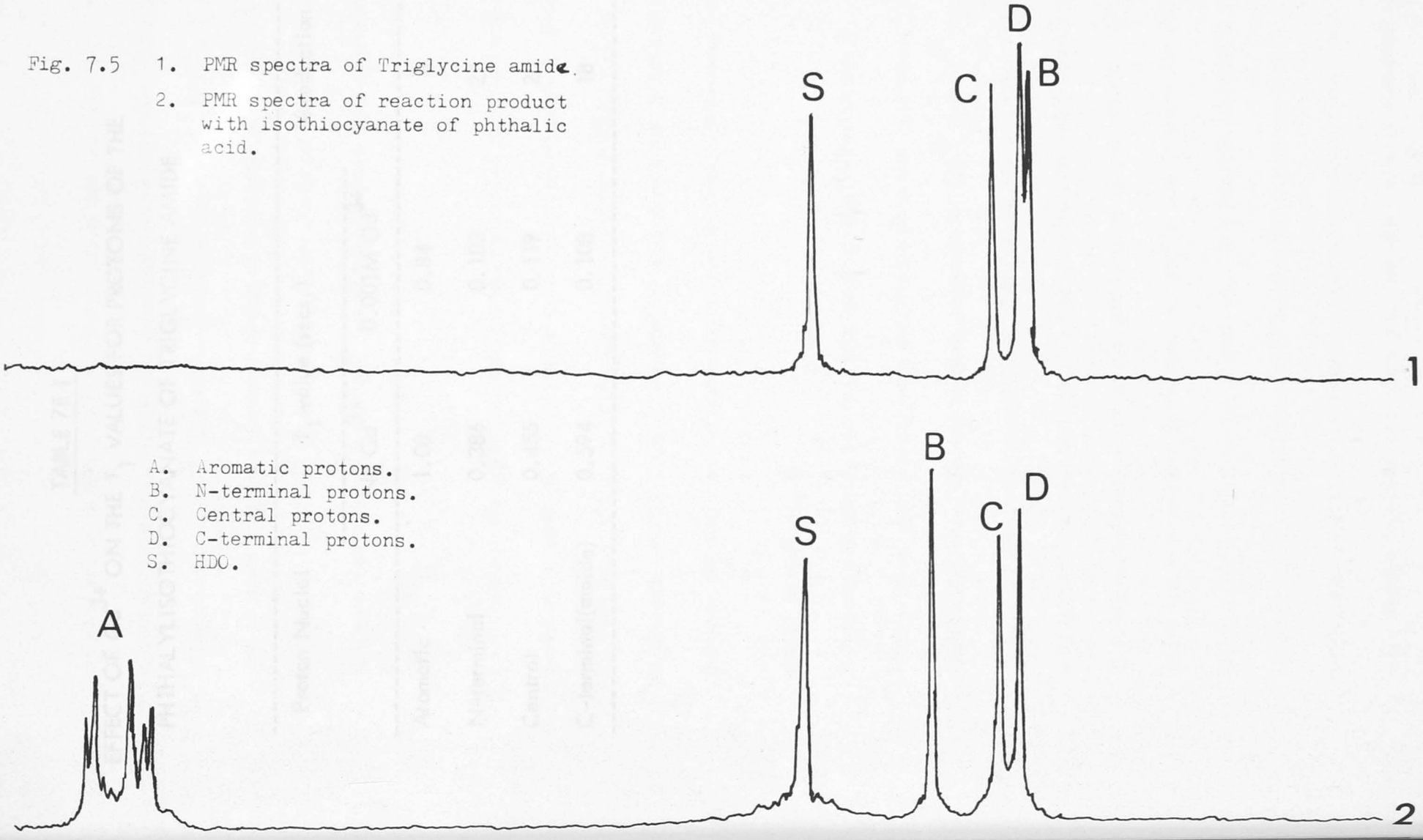


TABLE 7E I

EFFECT OF Gd^{3+} ON THE T_1 VALUES FOR PROTONS OF THE PHTHALYLISOTHIOCYANATE OF TRIGLYCINE AMIDE

Proton Nuclei	T_1 value (secs.)		% reduction
	No Gd^{3+}	0.001M Gd^{3+}	
Aromatic	1.08	0.84	8
N-terminal	0.386	0.103	27
Central	0.455	0.119	26
C-terminal(amide)	0.594	0.108	18

or more ligands bound by one or two carboxyl groups it is possible that some solution structure exists with similar distances between Gd ion and all α-proton nuclei resulting in the measured relaxation rates. At higher Gd^{3+} concentration more 1 : 1 complex would be expected. With a single ligand there should be greater freedom of movement in the peptide chain and sequential relaxation from the terminal site could be expected. It cannot be due to amide binding because T_1 measurements have already shown this to be negligible.

Because the isothiocyanate reaction can be carried out at pH 8.0 it is possible to use ester derivatives as carboxyl blocking groups. Reaction of (I) with diglycine methyl ester gave the expected product without any ester hydrolysis. Unfortunately this product also gave a precipitate with Pr^{3+} between pH 4 to 5 and it was not possible to measure a binding constant.

7H. Other Modifications

It has been recently reported (Marinetti et al, 1975) that the 3 nitrotyrosine moeity provides a specific binding site in proteins for lanthanide ions. The log binding constants for Gd^{3+} and Pr^{3+} were found to be 2.51 and 2.61 respectively. Indications are that such a binding site will provide a useful and specific site in certain proteins where nitration of tyrosine residues occurs at a single tyrosine (Stark, 1970). Although no experiments have been performed to study its use for sequence determination of peptides it seems worthy of mention. Any peptide containing a terminal tyrosine residue could presumably be nitrated to produce this binding site. With other peptides it would be necessary to attach a nitrotyrosine amino acid residue to either end of the peptide chain using a standard method.

CHAPTER 8

THE POTENTIAL OF NMR SPECTROSCOPY AS A SEQUENCING TECHNIQUE FOR PEPTIDES

Because of the low inherent sensitivity of NMR compared to other instrumental techniques it appears doubtful that it will ever rival sequencing methods currently in use such as the Edman method or mass spectrometry. Even with recent developments in NMR equipment, a thousand fold increase in sensitivity seems unlikely in the forseable future. However, there are many special cases where NMR spectroscopy and relaxation probes may prove extremely useful in providing sequence information.

8A. CHOICE OF NUCLEI

Both ^{1}H and ^{13}C nuclei have been studied with relaxation probes in the search for sequential magnetic perturbations. The general conclusions gained from this project are that ^{1}H nuclear resonances are preferable in that sensitivity is higher than ^{13}C (natural abundance), resulting in much less instrumentation time for spectral accumulation. Additionally, lower concentrations of paramagnetic ions are needed to cause relaxation of ^{1}H nuclei because the amount required is proportional to the gyro magnetic ratios of the respective nucleus. In effect this allows the relaxation effect to be sensed over a greater distance from site of binding before the point is reached where the binding site is saturated. This advantage is partly offset by the spin spin splittings which occur with ^{1}H nuclear resonances causing severe overlap. If sensitivity is not a problem and unlimited spectrometer time is available then ^{13}C sequencing has many advantages. Because of

the greater range of chemical shifts, there is less overlap of resonances, coupled with the advantage of singlet resonances from each nuclei. There is a further check on sequencing from both carbon nuclei of the peptide backbone and even the possibility of using the unique chemical shifts of carbonyl nuclear resonances which appear dependant on their nearest neighbour amino acid units.

8B. SEQUENCING FROM N OR C TERMINUS

NMR sequencing from the C-terminus of a peptide using Gd^{3+} is particularly attractive because of the lack of reliable C-terminal methods. Should the proposed T_1 difference spectra produce the expected results, then it may be possible to obtain the C-terminal sequence of a larger peptide or protein. To obtain this information by the Edman method would require the complete sequence of peptide from the N-terminus and with peptides in the region of 60 residues in length the Edman method would be reaching its limit in accuracy and length as the sequential degradation approached the C-terminus. N-terminal sequences of peptides can be obtained by NMR spectroscopy using Cu^{2+} for a direct method, or by a chemical modification procedure to attach Gd^{3+} binding sites to the N-terminus. It is, however, unlikely that this approach will compete with other N-terminal methods.

8C. SAMPLE SIZE

Samples in the picogram range have been sequenced by the Edman method. The NMR sequencing technique cannot approach these levels, but using microtubes may be adaptable to sub milligram amounts of most peptides. It is essential to maintain a high concentration of peptide to achieve maximum binding of paramagnetic ion, and the limiting factor in this case may then be peptide solubility, although the use of 6M GuDCl improves the solubility of peptides.

8D. SEQUENCING BY T_1 OR T_2 MEASUREMENT

For C-terminal sequencing with Gd^{3+} ions the measurement of selective reduction in T_1 values is preferable to measurement of T_2 values from resonance broadening experiments. T_1 values can be measured automatically with standard computer controlled spectrometers, and are affected at very low concentrations of Gd^{3+} which cause little or no resonance broadening. Hence resonances remain well resolved in T_1 experiments and do not broaden and merge together as in T_2 experiments.

For N-terminal sequence studies using Cu^{2+} probes it has not been possible to obtain reliable sequence information from T_1 experiments. In this case it is necessary to use T_2 measurements (sequential broadening) for the required information.

8E. CONCLUSIONS

To conclude, it appears that NMR spectroscopy with relaxation probes could usefully be applied for sequence determination of peptides from the N-terminus, using Cu^{2+}, and from the C-terminus using Gd^{3+}. N-terminal sequencing is not possible when histidine residues are present in the peptide except in the N-terminal position. In both cases sequence information of up to 5 residues from the site of binding of paramagnetic ion could be expected. However, in the presence of secondary binding sites for Gd^{3+}, i.e. side chain carboxyl groups, C-terminal sequencing is unsatisfactory. Unlike other techniques, NMR spectroscopy is a non-destructive technique and peptides sequenced by this method may be recovered intact.

REFERENCES

Abragam, J.W., (1961), Principles of nuclear magnetism. Oxford university Press, 527.

Agarwal, K.L., Jonhstone, R.A.W., Kenner, G.W., Millington, D.S., Sheppard, R.C., (1968), Nature, 219, 498.

Allerhand, A., Cochran, D.W., Doddrell, D., (1970), Proc. Natn. Acad. Sci (U.S.) 67, 1093.

Anteunis, M., Gelan, J., (1973), J. Amer.Chem.Soc., 95, 6502.

Anteunis, M., Callens, R., (1974), J.Mag.Res., 15, 317.

ApSimon, J.W., Beierbeck, H., (1972), J.C.S.Chem.Comm., 172.

Armitage, I., Dunsmore, G., Hall, L.D., Marshall, A.G., (1972), Can.J.Chem. 50, 2119.

Barry, C.D., North, A.C.T., Glasel, J.A., Williams, R.J.P., Xavier, A.V., (1971), Nature (London), 232, 236.

Beattie, J.K., Fensom, D.J., Freeman, H.C., (1976), J.Amer.Chem.Soc., 98, 500.

Bloembergen, N., (1957), J.Chem.Phys., 27, 572.

Bloembergen, J., Morgan, L.O., (1961), J.Chem.Phys. 34, 842.

Bloom, A.L., Schoolery, J.N., (1955), Phys.Rev., 97, 1251.

Bradbury, J.H., (1970), The physical principles and techniques of protein chemistry. Part B. Academic Press, New York, 99.

Bradbury, J.H., Crompton M.W., Warren, B., (1973), Anal.Biochem., 62, 310.

Campbell, I.D., Dobson, C.M., Williams, R.J.P., Xavier, A.V., (1973), J.Mag.Res., 11, 172.

Christl, M., Roberts, J.D., (1972), J.Amer.Chem.Soc., 94, 4565.

Cockerill, A.F., Rackham, D.M., (1970), Tett.Lett., 5149.

Conger, R.L., Selwood, P.W., (1952), J.Chem.Phys., 20, 383.

Connick, R.E., Fiat, D., (1966), J.Chem.Phys., 44, 4103.

Crompton, M., (1973), Honours Thesis, Australian National Univertity.

Cromwell, L.D., Stark, G.R., (1969), Biochem., 8.

Das, B.C., Gero, S.D., Lederer, E., (1967), Biochem.Biophys.Res.Comm., 29, 211.

Demarco, P.V., Elzey, T.K., Lewis, R.B., Wenkert, E., (1970), J.Amer.Chem.Soc., 92, 5734.

Dobbie, H., Kermack, W.D., (1955), Biochem. J., 59, 257.

Dowling, L.M., Stark, G.R., (1969), Biochem., 8.

Dwek, R.A., Richards, R.E., Morallee, K.G., Nieboer, E., Williams, R.J.P., Xavier, A.V., (1971), Eur, J.Biochem., 21, 204.

Dwek, R.A., (1973), in Nuclear magnetic resonance in biochemistry: Application to enzyme systems. Clarendon Press. Oxford.

Edman, P., (1950), Acta.Chem.Scan., 4, 277.

Edman, P., Begg, G., (1967), Eur.J.Biochem., 1, 80.

Einstein, A., (1906), Ann.Physik., 19, 289.

Emsley, J.W., Feeney, J., Sutcliffe, L.H., (1965), High resolution nuclear resonance spectroscopy. Pergamon Press. Oxford.

Espersen, W.G., Martin, R.B., (1976), J.Amer.Chem.Soc., 98, 40.

Frankel, L.S., (1969), J.Chem.Phys., 50, 943.

Freedman, M.H., Cohen, J.S., Chaiken, I.M., (1971), Biochem.Biophys.Res. Comm., 421148.

Freeman, H.C., Robinson, G., Schoone, J.C. (1964), Acta.Crystallogr., 17, 719.

Freeman, H.C., Schoone, J.C., Sime, J.G., (1965), Acta.Crystallogr., 18, 381.

Freeman, R., Hill, H.D.W., (1970), J.Chem.Phys., 53, 4103.

Freeman, R., Hill, H.D.W., (1971), J.Chem.Phys., 54, 3367.

Gould, D.C., Mason, H.S., (1973), Inorganic biochemistry. Elsevier, Amsterdam.

Graffeo, A.J., Bear, J.L., (1968), J.Inorg.Nucl.Chem., 30, 1577.

Gray, W.R., (1967), Methods Enzymol., 11, 469.

Hart, F.A., Moss, G.P., Staniforth, M.L., (1971), Tet.Let., 3389.

Hertz, J.E., Rodriguez, V.M., Hoseph-Nathan, P., (1971), Tet.Lett., 2949.

Hinckley, C.C., (1969), J.Amer.Chem.Soc., 91, 6160.

Horrocks, W.de W., (1970), Inorg.Chem., 9, 690.

Horsley, W.J., Sternlicht, H., (1968), J.Amer.Chem.Soc., 90, 3738.

Horsley, W.J., Cohen, J.S., (1970), J.Amer.Chem.Soc., 92, 680.

Howell, J., Private communication.

Huggins, M.L., (1942), J.Amer.Chem.Soc., 64, 2716.

Ingri, N., Sillen, L.G., (1964), Arkiv.Keni., 23, 47.

Jesson, J.P., (1967), J.Chem.Phys., 47, 579.

Karger, J., Pfeifer, H., (1968), Ann.des.Phys., 22, 51.

Kim, M.K., Martell, A.E., (1966), J.Amer.Chem.Soc., 88, 914.

Kim, M.K., Martell, A.E., (1969), J.Amer.Chem.Soc., 91, 872.

Kolat, R.S., Powell, J.E., (1962), Inorg.Chem., 1, 293.

Koltun, W.L., Roth, R.H., Gurd, F.R.N., (1963), J.Biol.Chem., 238, 124.

Konigsberg, W., (1967), Methods Enzymol, 11, 461.

Kurland, R.J., McGarvey, B.R., (1970), J.Mag.Res., 2, 286.

Korman, S., Clarke, H.T., (1956), J.Biol.Chem., 221, 113.

Latt, S.A., Vallee, B.L., (1971), Biochem., 10, 4254.

Laursen, R.A., (1966), J.Amer.Chem.Soc., 88, 5344.

Lewis, W.B., Jackson, J.A., Lemons, J.F., Taube, H., (1962), J.Chem.Phys., 36, 692.

Li, N.C., Scruggs, R.L., Becker, E.D., (1962), J.Amer.Chem.Soc., 84, 4650.

Li, N.C., Tang, P., (1964), J.Amer.Chem.Soc., 86, 1293.

Luz, Z., Meiboom, S., (1964), J.Chem.Phys., 40, 2686.

Maeda, H., Kawauchi, H., (1968), Biochem.Biophys.Res.Comm., 31, 188.

Marinetti, T.D., Snyder, G.H., Sykes, B.D., (1975), J.Amer.Chem.Soc., 97, 6562.

Markley, J.L., Horsley, W.J., Kelin, M.P., (1971), J.Chem.Phys., 55, 3604.

McConnell, H.M., Robertson, R.E., (1958), J.Chem.Phys., 29, 1361.

Morris, H.R., Williams, D.H., Ambler, R.P., (1971), Biochem.J., 125, 189.

Nakano, M., Nakano, N.I., Higuchi, T., (1967), J.Phys.Chem., 71, 3954.

Olsen, R.K., (1970), J.Org.Chem., 35, 1912.

Osterberg, R., Sjoberg, B., (1968), J.Biol.Chem., 243, 3038.

Peisach, J., Aisen, P., Blumberg, W.E., (1966), The Biochemistry of copper. Academic Press. New York.

Perrin, D.D., Agarwal, R.P., (1972), Pure & Appl.Chem., 34.

Pople, J.A., Schneider, W.G., Bernstein, H.J., (1959), High resolution nuclear magnetic resonance. McGraw Hill, New York.

Reubin, R., Fiat, D., (1967), Chem.Comm., 729.

Reubin, J., Reed, G.H., Cohen, M., (1970), J.Chem.Phys., 52, 161.

Reyes-Zamora, C., Tsai, C.S., (1971), Chem.Comm., 1047.

Robb, E.W., Choma, P.P., Onsager, E.K., (1972), Abs. 164th.Nat.Meet.Am. Chem.Soc. New York, No.ORGN-145.

Rondeau, R.E., Sievers, R.E., (1971), J.Amer.Chem.Soc., 93, 1522.

Rossotti, F.J.C., Rossotti, H.S., Whewell, R.C., (1971), J.Inorg.Nucl.Chem., 33, 2051.

Sabatini, A., Vacca, A., Gans, P., (1972), Talanta, 21, 53.

Sanders, J.K.M., Williams, D.H., (1971), J.Amer.Chem.Soc., 93, 641.

Sanders, J.K.M., Hanson, S.W., Williams, D.H., (1972), J.Amer.Chem.Soc., 94, 5325.

Sayce, I.G., (1968), Talanta, 15, 1397.

Schroeder, W.A., (1967), Methods Enzymol., 11, 445.

Sheinblatt, M., (1966a), J.Amer.Chem.Soc., 88, 2123.

Sheinblatt, M., (1966b), Ibid, 88, 2845.

Sheinblatt, M., (1967), Magnetic resonance in biological systems. Pergamon Press, Oxford.

Sherry, A.D., Yoshida, C., Birnbaum, E.R., Darnall, D.W., (1973), J.Amer.Chem.Soc. 95, 3011.

Shulman, R.G., Wyluda, B.J., (1959), J.Chem.Phys., 30, 335.

Silber, H.B., Farina, R.D., Swinehart, J.H., (1969), Inorg.Chem., 8, 819.

Solomon, I., (1955), Phys.Rev., 99, 559.

Spande, T.F., Witkop, B., Degani, Y., Patchornik, A., (1970), Adv.Prot. Chem., 24, 98.

Stark, G.R., (1968), Biochem., 7, 1796.

Sternlicht, H., (1965), J.Chem.Phys., 42, 2250.

Swift, T.J., Connick, R.E., (1962), J.Chem.Phys., 37, 307.

Sykes, B.D., (1969), J.Amer.Chem.Soc., 91, 949.

Sylva, R.N., Private communication.

Thomas, D.W., Das, B.C., Gero, S.D., Lederer, E., (1968a), Biochem. Biophys.Res.Comm., 32, 199.

Thomas, D.W., Das, B.C., Gero, S.D., Lederer, E., (1968b), Ibid, 32, 519.

Thomas, D.W., (1968c), Ibid, 33, 483.

Vliegenthart, J.F.G., Dorland, L., (1970), Biochem.J., 117, 31.

Vold, R.L., Waugh, J.S., Klein, M.P., Phelps D.E., (1968), J.Chem.Phys., 48, 3831.

Williams, R.J.P., (1970), Quart.Rev., 24, 331.

Woessner, D., (1962), J.Chem.Phys., 36, 1.

Woyski, M.M., Harris, R.E., (1963), In treatise on analytical chemistry, part II 8. Interscience. New York.

APPENDIX 1

(1) FOCAL PROGRAM for T_1 calculation from a standard inversion recovery experiment using a $(180\text{-}\tau\text{-}90)_n$ pulse sequence. This program performs a least squares analysis of $Ln(A_\infty - A_\tau)$ vs. τ where A_∞ and A_τ are the resonance peak heights in mm. for delay times of ∞ and τ resp. The program will ask for the peak height of the 90° pulse ($T=\infty$) followed by the number of delay times. Each delay time should then be entered with its corresponding peak height. For each delay time entered the value of $Ln(A_\infty - A_\tau)$ will be printed.

```
01.01  T  ! , "T1 CALC. FROM NMR INVERSION RECOVERY EXPT.
01.02  ERASE
01.03  A  ! , "90 DEGREE PULSE HEIGHT = "H0
01.04  A  ! , "NUMBER OF DELAY TIMES="N
01.05  T  %! , "DELAY(MS)     HEIGHT (MM)     LOS (H0-H(I))"
01.06  F  I-1, N;D0 6
01.20  F  I=1, N;D0 4

02.08  S  M=((N*J)-(K*L))/((N*P)-(K ↑ 2))
02.18  S  B=((P*L-(K*J))/((N*P)-K ↑ 2))
```

```
03.08   T   ! ! ;S TT=-(1.0/M)
03.18   T   %,"SPIN LATTICE RELAXATION TIME=TT,"MILLISECS"
03.20   F   I=1,N;S CA(I)=M*X(I)+B
03.22   S   DF=O
03.24   F   I=1,N;S DF=DF+(CA(I)-Y(I))↑2
03.26   S   SD=FSQT(DF/(N-1))
03.28   T   !,"STD.DEV.="SD
03.30   Q


04.10   S   K=K+X(I);S L=L+Y(I)
04.20   S   P=P+(X(I))↑2;S J=J+(X(I)*Y(I))


06.02   A   !,X(I),"                ",H(I),"                "
06.03   S   Y(I)=FLOG(HO-H(I));T Y(I)
```

(2) FOCAL PROGRAM for theoretical T_1 curves for overlapping resonances where each nuclei has a different T_1 value.

```
01.01  ERASE
01.02  S  U=1
01.03  T  !,"THEORETICAL T1 CURVES FOR OVERLAPPING PEAKS" !
01.04  A  !,"NO. OF EXPERIMENTS"X
01.05  A  !,"NO. OF DELAY TIMES"N
01.07  F  I=1,N;A  ."DELAY="D(I)
01.08  T  !,"**************************************************
01.09  T  %1.00  ,"EXPERIMENT NO."U
01.10  A  !,"FIRST T1(SECS)="T1;A, "SECOND T1(SECS.)="T2
01.12  S  MT=20.0;S MA=10.0;S MB=10.0
01.15  F  I=1,N;S W(I)=MA*(1.0-2*FEXP(-(D(I)/T1)))
01.17  F  I=1,N;S X(I)-MB* (1.0-2*FEXP(-(D(I)/T2)))
01.20  F  I=1,N;S X(I)=W(I)+X(I)
01.22  F  I=1,N;S R(I)=FLOG(MT-Z(I))
01.25  T  !,"DELAY(SECS.)  PEAK HEIGHT(MM)      LN VALUE"
01.26  T  %6.03  !,"     0.000       "-(MT),"       ", FLOG(2*MT)
01.28  F  I=1,N;T  ,D(I),"          ",Z(I),"          ",R(I)
01.30  S  U=1+U
01.32  I  (U-1-X)1.08,1.36,1.36
01.36  Q
```

(3) FOCAL PROGRAM for theoretical T_1 difference curves. The T_1 values expected in the absence and presence of Gd^{3+} are entered with the desired delay time (τ). The program will give the difference in peak heights between the samples with and without Gd^{3+} at the same τ value.

```
01.01  ERASE
01.02  S  U=1
01.03  T  !, "THEORETICAL CURVES FOR T1 DIFF. SPECTRA",
01.04  A  !, "NUMBER OF EXPTS."X
01.05  A  !, "NUMBER OF DELAY TIMES"N
01.07  F  I=1,N;A  ,"DELAY="D(I)
01.08  T  !!,"*****************************************
01.09  T  %1.00  ,"EXPERIMENT NO. "U
01.10  A  !," T1(NO GD.)="T1;A ,"T1(+GD.)="T2
01.12  S  MA=10.0;S MB=10.0
01.15  F  I=1,N;S W(I)=MA*(1.0-FEXP -(D(I)/T1)))
01.17  F  I=1,N;S X(I)=MB*(1.0-2*FEXP(-(D(I)/T2)))
01.20  F  I=1,N;S Z(I)=W(I)-X(I)
01.25  T  !  ,"DELAY(SECS.)   PEAK HEIGHT(MM)
01.28  F  I=1,N;T %6.03  , D(I)," ",Z(I)
01.30  S  U=1+U
01.32  I  (U-1-X)1.08,1.36,1.36
01.36  Q
```

(4) FOCAL PROGRAM for theoretical binding curves for 1 : 1 complexes of metal and ligand with binding constants.

```
01.02  T  "THEORETICAL BINDING CURVES FOR 1 : 1 COMPLEXES", !
01.04  A  "NO. VALUES K="N, !
01.06  F  X=1, N;A "K="K(X), !
01.08  A  "NO. SUBSTRATE CONCS.="I,
01.10  F  Y=1, I;A "S="S(Y), !
01.14  A  "INITIAL METAL CONC. ="R, !
01.16  A  "FINAL METAL CONC. =R2, !
01.18  A  "INCREMENT IN METAL CONC.="Z, !

02.02  F  Y=1, I;DO 3
02.03  Q

03.01  T  ! ! ,T %2.00,"CURVE "Y,  ;T %,"SUBSTRATE CONC.="S(Y),
03.02  T  ! !,"METAL          CONC.OF COMPLEXED METAL",  ;S R1=R
03.03  T  "      ",F X=1, N;T %9.02, K(X)
03.04  T  !;T %5.04, R1;F X=1, N;DO 4
03.05  S  R1=R1*Z
03.08  I  (R1-R2)3.04, 3.04, 3.10
03.10  T  !;RETURN
```

```
04.08  S  B=(S(Y)+(R1)+1/K(X))
04.10  S  C=S(Y)*R1
04.12  S  A=(B-FSQT(B  2-4*C))/2
04.13  I  (A)4.14,4.14,4.16
04.14  S  A=(B+FSQT(B  2-4*C))/2
04.16  T  %8.05,A
04.18  RETURN
```

(5) FOCAL PROGRAM for calculation of binding constants by proton relaxation enhancement (PRE).

```
01.02  A  !,"N="N
01.05  F  I=1,N;A  ,"L="L(I),"T1="T(I)
01.07  A  !,"HDO="H,"HDO+GD="G,  "MT="M
01.10  F  I=1,N;S R(I)=(1/T(I)-L/H)/(1/G-1/H)
01.14  F  I=1,N;S X(I)=1/L(I)
01.15  I  (FABS(K-KT-0.1))2.30,2.30,1.16
01.16  F  I=1,N;S Y(I)=1/R(I)
01.17  S  KT=K
01.18  S  K=0;S L=0;S P=0;S J=0
01.19  F  I=1,N;DO 6
01.20  DO 7
01.25  S  K=S/I;T %  ,"KD="K;S D=1+1/I:T  ,"EB="D

02.02  F  I=1,N;DO 8
02.12  F  I=1,N;S W(I)=(XX(I)*M)/K+XX(I))
02.15  F  I=1,N;S R(I)=((W(I)*D)/M)+((M-W(I))/M)
02.20  GOTO 1.15
02.30  Q

06.01  S  K=K+X(I);S L=L+Y(I)
06.02  S  P=P+(X(I))'2;S J=J+(X(I)*Y(I))
```

```
07.02   S   S=((N*J)-(K*L))/((N*P)-(K'2))

07.04   S   I=((P*L)-(K*J))/((N*P)-(K'2))


08.04   S   XX(I)=((M+L(I)-K)-FSQT((M+L(I)-K)'2-4*K*L(I)))*0.5

08.06   I   (XX(I)) 8.08,8.08,8.09

08.07   I   ((XX(I)/L(I))-1) 8.09,8.09,8.08

08.08   S   XX(I)=((M+L(I)-K)+FSQT((M+L(I)-K)'2-4*K*L(I)))*0.5

08.09   RETURN
```

(6) FOCAL PROGRAM for calculation of pK values from NMR shift data.

```
01.10  A  "DA="DA
01.11  A  "NO. POINTS",N
01.13  F  Z=1,N;A "D(Z)=",D(Z)
01.15  F  Z=1,N;A "PH(Z)=",PH(Z)

02.02  F  Z=1,N;S PH(Z)=FEXP(-PH(Z)*2.3025)
02.20  S  SX=O;S SD=O;S QX=O;S DX=O
02.21  F  Z=1,N;S SX=SX+PH(Z)*(D(X)-DA)
02.22  F  Z=1,N;S SD=SD+D(Z)
02.23  F  Z=1,N;S QX=QX+(PH(X)*(D(Z)-DA))↑2
02.24  F  Z=1,N;S DX=DX+PH(Z)*(D(Z)-DA)*D(Z)
02.40  S  M=(N*DX-SX*SD)/(N*QX-SX↑2);S C=(QX*SD-SX*DX)/
          (N*QX-SX↑2)

03.01  F  Z=1,N;S Y(Z)=(M*PH(Z)*(D(Z)-DA)+C-D(Z))↑2
03.02  S  SC=O
03.03  F  Z=1,N;S SC=SC+Y(Z)
03.04  T  !!,%,"C=",C,"        SC=",FSQT(SC/N-1)),
03.11  S  SM=FSQT((SC/N-1))*(N*QX-SX↑2)));S SM=SM/M↑2
```

```
04.01   S   M=1/M;S PK=-FLOG(M)/2.3025
04.03   S   M=(FLOG(M)/2.3025)-(FLOG(M+SM)/2.3025);T "PK=", PK, "
            +OR-", M
04.04   T   ! !
04.05   F   Z=1, N;T %, (Z)*(D(Z)-DA), !
04.06   Q
```

(7) FORTRAN PROGRAM BBW651*SHIFT.SANDERS for binding constants from NMR shift data.

```
C  EQUIL CONSTANTS FROM NMR SHIFT DATA
C  REF. METHOD OF WILLIAMS ET AL J.A.C.S.(1972)94, 5325
C  SO=SUBSTRATE, LO=LANTHANIDE, DELTA=SHIFT(PPM)
      DIMENSION SO(50), LO(50), DELTA(50), X(50), Y(50), CALCY(50)
      DIMENSION TITLE(7)
      REAL LO, INCR, LAST, INTER, K, L, J, N, KB
      INTEGER TITLE
      READ(5, 100)TITLE
  100 FORMAT(7A6)
      READ(5, 101)NUM
  101 FORMAT(12)
      N=NUM
      READ(5, 103)FIRST, INCR, LAST
      DO 2 I=1, NUM
      READ(5, 103)LO(I), SO(I), DELTA(I)
    2 CONTINUE
  103 FORMAT(3E14.7)
      DO 7 I=1.NUM
      IF(FIRST.LE.DELTA(I))FIRST=DELTA(I)+INCR
    7 CONTINUE
      TEST=0.0
```

```
 8 DO 3 I=1, NUM
   Y(I)=(LO(I)*FIRST)/(SO(I)*DELTA(I))
   X(I)=1.0/(SO(I)*(1.0-(DELTA(I)/FIRST)))
 3 CONTINUE
   L=0
   P=0
   J=0
   DO 14 I=1, NUM
   K=K+X(I)
   L=L+Y(I)
   P=P+(X(I))**2
   J=J+(X(I)*Y(I))
14 CONTINUE
   SLOPE=((N*J)-(K*L))/((N*P)-(K**2))
   INTER=((P*L)-(K*J))/((N*P)-(K**2))
   DO 4 I=1, NUM
   CALCY(I)=SLOPE*X(I)+INTER
 4 CONTINUE
   DEV=0.0
   DO 5 I=1, NUM
   DEV=DEV+(CALCY(I)-Y(I))**2
 5 CONTINUE
   IF(TEST.LE.0.0) GO TO 6
   IF(DEV.GT.TEST) GO TO 10
```

```
  6 TEST=DEV
    FIRST-FIRST+INCR
    IF(FIRST.LT.LAST) GO TO 8
 10 WRITE(6.106) TITLE
    WRITE(6, 207)
207 FORMAT(1X, 'PROGRAM BBW651*SHIFT.SANDERS')
106 FORMAT(1X,7A6)
    KB=1.0/SLOPE
    WRITE(6, 107)KB, FIRST
107 FORMAT(1X, 'EQUILIBRIUM CONSTANT=', E14.7/1X.
   +'MAXIMUM SHIFT=', E14.7, 'P.P.M.')
    SIGMA=DEV/(N-1)
    WRITE(6, 108)SIGMA
108 FORMAT(1X, 'SIGMA=', E14.7)
    WRITE(6, 118)
110 FORMAT(1X, 'SUBSTRATE', 1X, 'LANTHANIDE', 1X, 'DELTA(PPM)',
    1X, 'X(I)')
    DO 99 I=1, NUM
 99 WRITE(6, 111)SO(I), LO(I), DELTA(I), X(I)
111 FORMAT(1X, 4E14.7)
    STOP
    END
```

(8) FORTRAN PROGRAM BBW651*SHIFT.SHERRY for binding constants from NMR shift data.

```
C  EQUILIBRIUM CONSTANTS FROM SHIFT DATA
C  REF. SHERRY ET. AL. J.A.C.S.95,3011,(1973)
C  FOR UP TO 50DATA POINTS
      DIMENSION SO(50),LO(50),DELTA(50),X(50),Y(50),
      +ZLO(50),TITLE(50)
      REAL SO,LO,INTER,K,LAST,INCR,N
   90 FORMAT(7A6)
  100 FORMAT(I2)
      READ(5,90)TITLE
      READ(5,100)NUM
      READ(5,250)FIRST,INCR,LAST
      DO 2 I=1,NUM
      READ(5,250)LO(I),SO(I),DELTA(I)
    2 CONTINUE
  250 FORMAT(3E14.7)
      N=NUM
      DO 29 I=1,NUM
   29 ZLO(I)=LO(I)
      K=0.0
    5 TEST=K
```

```
      DO 3 I=1,NUM
      Y(I)=1.0/DELTA(I)
      X(I)=1.0/LO(I)
    3 CONTINUE
      XK=0
      XL=0
      P=0
      XJ=0
      DO 14 I=1,NUM
      XK=XK+X(I)
      XL=XL+Y(I)
      P=P+(X(I))**2
      XJ=XJ+(X(I)*Y(I))
   14 CONTINUE
      SLOPE=((N*XY)-(XK*XL))/((N*P)-(XK**2))
      INTER=((P*XL)-(XK*XJ))/((N*P)-(XK**2))
      HZERO=1.0/INTER
      K=1.0/(SLOPE*HZERO)
      DO 4 I=1,NUM
    4 LO(I)=DELTA(I)/(K*(HZERO-DELTA(I)))
      IF((K-TEST).GE.0.1)GO TO 5
      WRITE(6,95)TITLE
   95 FORMAT(1X,7A6)
```

```
103 FORMAT(1X,'EQUILIBRIUM CONSTANT=',E14.7/1X,'MAX.
    SHIFT=',E14.7)
    WRITE(6,400)
400 FORMAT(1X,'INIT.LANTHANIDE',1X,'equil.lanthanide',1X,'LIGAND.
    +CONC.',4X,'X(I)',8X,'Y(I)',4X,'DELTA')
    DO 7 I=1,NUM
    WRITE(6,300)ZLO(I),LO(I),SO(I),X(I),Y(I),DELTA(I)
  7 CONTINUE
300 FORMAT(1X,6E14.7)
    WRITE(6,103)K,HZERO
    STOP
    END
```

(9) FORTRAN PROGRAM BBW651*SHIFT.SYKES for binding constants from NMR shift data.

```
C  EQUIL. CONST. FROM NMR SHIFT DATA (AT CONST. LANTHANIDE)
C  REF. SYKES J.A.C.S.81,949,(1969)
C  PROGRAM FOR UP TO 50 DATA POINTS
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION LO(50),SO(50),DELTA(50),LS(50),X(50),Y(50),W(50)
     +,CALCY(50),TITLE(50)
      REAL LO,INTER,LS,LAST,INCR,N,K,L,J
      INTEGER TITLE
      READ(5,100)TITLE
  100 FORMAT(7A6)
      READ(5,200)NUM
      N=NUM
  200 FORMAT(I2)
      READ(5,300)FIRST,INCR,LAST
  300 FORMAT(3E14.7)
      DO 2 I=1,NUM
      READ(5,300)LO(I),SO(I),DELTA(I)
    2 CONTINUE
      TEST=0.0
   20 DO I=1,NUM
      W(I)=LO(I)+FIRST+SO(I)
```

```
 3 CONTINUE
   DO 4 I=1, NUM
   LS(I)=W(I)-DSQRT((W(I))**2-(4.0*LO(I)*SO(I)))/2.0
   IF(LS(I)/LO(I).LT.O.O.OR.LS(I)/LO(I).GT.1.0)
  +LS(I)=(W(I)+DSQRT(W(I)**2-(4.0*LO(I)*SO(I)))/2.0
 4 CONTINUE
   DO 5 I=1, NUM
   Y(I)=DELTA(I)
   X(I)=LS(I)/SO(I)
 5 CONTINUE
   K=0
   L=0
   J=0
   P=0
   DO 6 I=1, NUM
   K=K+X(I)
   L=L+Y(I)
   J=J+(X(I)*Y(I))
 6 CONTINUE
   SLOPE=((N*J)-(K*L))/((N*P)-(K**2))
   INTER=((P*L)-(K*J))/((N*P)-(K**2))
   DO 7 I=1, NUM
   CALCY(I)=SLOPE*X(I)+INTER
```

```
    7 CONTINUE
      DEV=0.0
      DO 8 I=1,NUM
      DVE=DEV=(Y(I)-CALCY(I))**2
    8 CONTINUE
      IF(TEST.LE.0.0) GO TO 9
      IF(DEV.GT.TEST) GO TO 11
    9 TEST=DEV
      FIRST=FIRST+INCR
      IF(FIRST.LE.LAST) TO TO 20
   11 SIGMA=DSQRT(DEV/(N-1))
      WRITE(6,400)TITLE
  400 FORMAT(1X,'PROGRAM BBW651*SHIFT.SYKES',/1X,7A6)
      WRITE(6,500)FIRST,SLOPE
  500 FORMAT(1X,'EQUIL.CONST.=',E14.7,/1X,'MAX SHIFT=',E14.7)
      WRITE(6,600)
  600 FORMAT(1X,'SUBSTRATE',5X,'LANTHANIDE',4X,'COMPLEX',
     +7X,'DELTA',9X,'X(I)',9X,'Y(I)')
      DO 10 I=1,NUM
      WRITE(6,700)SO(I),LO(I),LS(I),DELTA(I),X(I),Y(I)
  700 FORMAT(1X,6E14.7)
   10 CONTINUE
      STOP
      END
```

(10) FORTRAN PROGRAM BBW651*SHIFT.NAKANO for binding constants from NMR shift data.

```
C  EQUILIBRIUM CONSTANTS FROM NMR SHIFT DATA
C  REF. NAKANO ET AL. J.PHYS.CHEM.71,3954,(1967)
C  FOR UP TO 50 DATA POINTS
      DIMENSION LO(50,SO(50),DELTA(50),X(50),Y(50),LS(50)
      DIMENSION TITLE(7)
      REAL LO,LS,J,L,K,N,INTER,MAX,KB,LAST,INCR
      READ(5,100)TITLE
  100 FORMAT(7A6)
      READ(5,101)NUM
  101 FORMAT(I2)
      SLOPE=0.0
      TEST=0.0
      READ(5,102)FIRST,INCR,LAST
      N=NUM
      DO 2 I=1,NUM
      READ(5,102)LO(I),SO(I),DELTA(I)
    2 CONTINUE
  102 FORMAT(3E14.7)
      DO 6 I=1,NUM
    6 LS(I)=0.0
   10 TEST=SLOPE
```

```
      DO 6 I=1, NUM
    6 LS(I)=0.0
   10 TEST=SLOPE
      DO 3 I=1, NUM
      Y(I)=LO(I)/DELTA(I)
      X(I)=(LO(I)+SO(I)-LS(I))
    3 CONTINUE
      K=0
      L=0
      P=0
      J=0
      DO 14 I=1, NUM
      K=K+X(I)
      L=L+Y(I)
      P=P+(X(I))**2
      J=J+(X(I)*Y(I))
   14 CONTINUE
      SLOPE=((N*J)-(K*L))/((N*P)-(K**2))
      INTER=((P*L)-(K*J))/((N*P)-(K**2))
      DO 4 I=1, NUM
      LS(I)=DELTA(I)*SO(I)*SLOPE
    4 CONTINUE
      IF((SLOPE-TEST).GE.0.001) GO TO 10
      MAX=1.0/SLOPE
      KB=SLOPE/INTER
```

```
      WRITE(6,104)TITLE
  104 FORMAT(1X,'PROGRAM BBW651*SHIFT.NAKANO',/1X,20A4)
      WRITE(6,105)KB,MAX
  105 FORMAT(1X,'EQUILIBRIUM CONST.='E14.7/1X,'MAX SHIFT=',E14.7)
      WRITE(6,105)
  106 FORMAT(1X,'SUBSTRATE',2X,'LANTHANIDE',2X,'COMPLEX CONC.'
     +,1X,'X(I)',7X,'Y(I)',7X,'DELTA')
      DO 12 I=1,NUM
   12 WRITE(6,107)SO(I),LO(I),LS(I),X(I),Y(I),DELTA(I)
  107 FORMAT(1X,6E14.7)
      STOP
      END
```

(11) FORTRAN PROGRAM SIAS for species distribution of multi equilibria in solution. This listing is included because it is an unpublished program (Sylva, 1974)

```
C PROGRAM SIAS                    IMPUT DATA
C NUMBER OF SETS OF EXPERIMENTAL CONDITIONS TO BE RUN (12)
C TITLE OF FIRST (SECOND THIRD ETC) EXPERIMENT: ANY
C CHARACTERS IN COLS. 1-80
C SERIES OF 5CARDS(5IS MANDATORY) COMPRISING THE ARRAY TEXT
C WHEREIN THE USER PROVIDES A DESCRIPTION OF THE METALS AND
C LIGANDS FOR RECORDING IN THE OUTPUT: ANY CHARACTERS INCLUDING
C BLANKS IN COLS. 1-80
C NUMBER OF LIGANDS  , OF METALS AND COMPLEX SPECIES FORMED
C (INCLUDING PROTONATED LIGANDS AND HYDROLYSED METAL IONS
C (212,13)
C SPECIES NO AND SPECIES DESCRIPTION 1-14. THE NO OF MOLECULES
C OF LIGAND (1),(2),(3) etc. UP TO 10. THE NO OF METAL IONS:
C (1),(2),(3),ETC. UP TO 10. THE NO OF HYDROXYL IONS (A POSITIVE
C INTEGER), OR OF PROTONS(A NEGATIVE INTEGER). THE LOG OF THE
C CUMULATIVE FORMATION CONSTANTS OF THE SPECIES EXCEPT FOR
C HYDROLYSED SPECIES WHERE LOG OF THE CUMULATIVE HYDROLYSIS
C CONSTANTS, *BETAN? IS USED.  INDEX TO SPECIES SCANS DESCRIBED
C BY METAL NO. (1)-1,(2)-2,(3)-3 ETC. FOR LIGAND ONLY INDEX IS
```

```
C LEFT BLANK.(15A1,5X,2I12,6X,F8.4,I2)
C THE TOTAL CONC. OF EACH LIGAND.(8E10.3)
C THE TOTAL CONC. OF EACH METAL ION. (8E10.3)
C A SERIES OF CARDS BEARING PH AND INDEX. INDEX=OOR BLANK
C FOR ALL BUT LAST CARD OF EXPERIMENT WHEN INDEX =1.(F10.4,I1)
C THEN RETURN TO ITEM 2 UNTIL THE TOTAL NUMBER OF EXPERIMENTS
C GIVEN INITIALLY HAS BEEN REACHED.
C ***************************************************************
C NOTE
C CHANGE OF PH IS SHOWN ON PH CARDS AND IS HANDLED WITHIN A
C SINGLE EXPERIMENT
C CHANGE OF LIGAND OR ITS CONCENTRATION, METAL ION OR ITS
C CONCENTRATION REQUIRES DIFFERENT EXPERIMENT FOR EACH
C SUCH CHANGE AND A FULL DATA DECK FOR EACH EXPERIMENT
C OH IS NOT TREATED AS A LIGAND BUT THE PROGRAM HAS BUILT
C IN FEATURES WHICH REQUIRE THE USE OF THE HYDROLYSIS
C CONSTANTS FOR THE HYDROLYSED SPECIES.
C EG. *BETAN=(M(OH)N)(H+)EXPN/(M+).OBTAINED FROM BETAN
C VALUES - (M(OH)N)/(M+)(OHEXPN FROM THE RELATIONSHIP:
C LOG*BETAN=LOGBETAN+N*14
      IMPLICIT REAL*(A-H,O-Z)
      REAL*4 TITLE,TEXT
      DIMENSION E(49),DM(10,DMY(10),TITLE(20),TEXT(20,5),PFM(10)
     +,PCS(48),IML(48),SNAME(48,15)
      COMMON C(49),Y1(10),Y2(10),Y3(10),Y4(10),BTOT(10),CLTOT(10),
```

```
     +TX(10), VX(10), ML(10, 48)MN(49)AL(10, 48)AM(10, 48), MM(10, 48),
     +AN(48), B(48)
      COMMON NL, NM, N, IPT, UK
    1 FORMAT(I2)
    2 FORMAT(15A1, 5X, 2I2, 6X, F8.4, I2)
    6 FORMAT(1X, 15A1, 2X, 20(2X, I2), 3X, I2, 2X, F8.4)
    8 FORMAT(8E10.3)
    9 FORMAT(22H TOTAL CONC. OF METAL (, I2, 4H)  =, D10.3)
   10 FORMAT(25H TOTAL CONCN. OF LIGAND(, I2, 4H)    =, D10.3)
   11 FORMAT(23H, 'C1          C2          C3          C4          C5
     +C6          C7          C8          C9          C10')
   12 FORMAT(2I2, I3)
   17 FORMAT(F10.4, I1)
   30 FORMAT(1HO, 'PH= ', F6.3)
   33 FORMAT(1X, 20X, 'L1 Ls L3 L4 L5 L6 L7 L8 L9 L10 M1 M2 M3
      + M4 M5 M6 M7 M8 M9 M10 OH LOG BETA')
   34 FORMAT(1HO, 'FREE METALS', 6X, 10(1X, 1PE10.3))
   36 FORMAT(1HO, 'FREE LIGANDS', 4X, 10(1X, 1PE10.3))
   37 FORMAT(1HO, 'COMPLEX SPECIES')
   38 FORMAT(1HO, '% COMPLEX SPECIES')
   43 FORMAT(1X, L3, '-', L3, 10X, 10(1X, 1PE10.3))
   53 FORMAT(1X, I3, '-', I3, 10X, 10(1X, 1PE10.3))
   63 FORMAT(1HO, 13X, '   DISTRIBUTION OF METAL', I3, 28X, '%', 12X,
      'CONC')
   73 FORMAT(40X, 15A1, I3, 2(5X, 1PE10.3))
```

```
 83 FORMAT(1HO,57X,'TOTAL',1PE10.3,5X,1PE10.3)
 93 FORMAT(52X,'FREEM',5X,1PE10.3,5X,1PE10.3)
 90 FORMAT(64X,'------------',6X,'------------')
122 FORMAT(/1HO)
123 FORMAT(20A4)
124 FORMAT(1X,20A4/)
125 FORMAT(1HO,'% FREE METALS',4X,10(1X,1PE10.3))
130 FORMAT(20A4)
131 FORMAT(1X,20A4//)
234 FORMAT(1HO)
800 FORMAT(1HO,'**************************************************
   +************************************************************
   +**************')
900 FORMAT (1H1)
    READ (5,1)NJ
    NJD=0
106 WRITE(6,900)
    READ(5,130)TITLE
    WRITE(6,131)(TITLE(I), I=1,20)
    WRITE(6,33)
    READ(5,123)TEXT
    READ(5,12)NL,NM,N
    DO 7 J=1,N
  3 READ(5,2)(SNAME(J,I),I=1,15),(ML(I,J),I=1,10),(MM(I,J),I=1,10),MN
   +(J),E(J),IML(J)
```

```
 7 WRITE(6,6)(SNAME(J,I),I=1,15),(ML(I,J),I=1,10),(MM(I,J),I=1,10),
  +MN(J),E(J)
    WRITE(6,122)
    WRITE(6,124)TEXT
    READ(5,8)(CLTOT(I),I=1,NL)
    READ(5,8)(BOTO(I),I=1,NM)
    WRITE(6,10)(I,CLTOT(I),I=1,NL)
    WRITE(6,9)(i,BTOT(I),I=1,NM)
    WRITE(6,900)
    HX=DLOG(10.DO)
    IPT=1
    DO 4 J=1,N
    AN(J)=MN(J)
    DO 4 I=1,10
    AL(I,J)=ML(I,J,)
 4  AM(I,J)=MM(I,J)
    DO 13 I=1,N
13  B(I)=DEXP(HX*E(I))
    DO 14 I=1,NM
14  Y1(I)=BTOT(I)*1.D-005
    DO 15 I=1,NL
15  Y3(I)=CLTOT(I)*1.D-005
16  READ(5,17)PH,INDEX
    WRITE(6,30)PH
```

```
      WRITE(6,234)
      UX=DEXP(HX*PH)
      IF(IPT-1)18,18,27
   18 DO 19 I=1,NM
   19 VX(I)=BTOT(I)
      DO 20 I=1,NL
      DMY(I)=1.DO
      DO 22 J=1,N
      IF(ML(I,J))22,22,200
  200 DO 21 K=1,NM
      IF(MM(K,J))22,21,22
   21 CONTINUE
      DM(I)=(DEXP(HX*E(J)))*UX**MN(J)
      MNY(I)=DMY(I)+DM(I)
   22 CONTINUE
   20 CONTINUED
      DO 23 I=1,NL
   23 TX(I)=CLTOT(I)/DMY(I)
   27 CALL COGS
      WRITE(6,11)
      WRITE(6,34)(VZ(I),I=1,NM)
      WRITE(6,36(TX(I),I=1,NL)
      DO 500 I=1,NM
  500 PFM(I)=100.DO*VX(I)/BTOT(I)
```

```
      DO 550 J=1,N
      MC=IML(J)
      IF(MC.GT.10) GO TO 520
      PCS(J)=100.DO*C(J)*MM(MC,J)/BTOT(MC)
      GO TO 550
  520 LC=IML(J)-10
      PCS(J)=100.DO*C(J)/CLTOT(LC)
  550 CONTINUE
      WRITE(6,125)(PFM(I),I=1,NM)
      DO 350 I=1,NM
      WRITE(6,_63)I
      TC=0.0
      TP=0.0
      DO 300 J=1,N
      MC=IML(J)
      IF(MC.NE.I) GO TO 300
      WRITE(6,73) (SNAME(J,NB),NB=1,15),J,PCS(J),C(J)
      TP=TP+PCS(J)
      TC=TC+C(J)
  300 CONTINUE
      TTP=TP+PFM(I)
      TTC=TC+VX(I)
      WRITE(6,93)PFM(I),VX(I)
      WRITE(6,90)
      WRITE(6,83)TTP,TTC
```

```
350 CONTINUE
    WRITE(6,234)
    WRITE(6,800)
    IF(INDEX)31,16,31
 31 NJD=NJD+1
    IF(NJD-NJ)106,32,32
 32 STOP
    END
```

```
      SUBROUTINE COGS
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION TERM(48),TERN(48),ALO(10),BO(10),TY(10),VY(10)
      COMMON C(48),Y1(10),Y2(10),Y3(10),Y4(10),BTOT(10),CLTOT(10),
     +TX(10),VX(10),ML(10,48),MN(48),AL(10,48),AM(10,48),MM(10,48),
     +AN(48),B(48)
      COMMON NL,NM,N,IPT,UX
   99 FORMAT(1HO,'NUMBER OF ITERATIONS',I4)
  234 FORMAT(1HO)
  998 FORMAT(1HO,'ITERATION DID NOT CONVERGE')
      NIT=0
      DO 20 K=1,N
   20 TERM(K)=B(K)*UX**MN(K)
    2 DO 3 K=1,N
    3 TERN(K)=TERM(K)
      DO 21 K=1,N
      DO 23 J=1,NM
   23 TERN(K)=TERN(K)*VX(J)**NM(j,k)
   21 CONTINUE
      DO 22 K=1,N
      DO 15 J=1,NL
   15 TERN(K)=TERN(K)*TX(J)**ML(J,K)
    5 C(K)=TERN(K)
   22 CONTINUE
      NIT=NIT+1
      DO 7 I=1,NM
```

```
      BO(I)=VX(I)
      DO 8 K=1,N
    8 BO(I)=BO(I)+AM(I,K)*C(K)
      RATIO=BO(I)/BTOT(I)
      VY(I)=VX(I)/DSQRT(RATIO)
    7 Y2(I)=DABS(BO(I)-BTOT(I))
      DO 9 I=1,NL
      ALO(I)=TX(I)
      DO 10 K=1,N
   10 ALO(I)=ALO(I)+AL(I,K)*C(K)
      RATIO=ALO(I)/CLTOT(I)
      TY(I)=TX(I)/DSQRT(RATIO)
    9 Y4+DABS(ALO(I)-CLTOT(I))
      IF(NIT-999)11,11,999
   11 DO 12 I=1,NM
      IF(Y1(I)-Y2(I))14,12,12
   12 CONTINUE
      DO 13 I=1,NL
      IF(Y3(I)-Y4(I))14,13,13
   13 CONTINUE
      IPT=IPT+1
      WRITE(6,99)NIT
      WRITE(6,234)
      RETURN
   14 DO 16 I=1,NL
```

```
   16 TX(I)=TY(I)
      DO 17 I=1,NM
   17 VX(I)=VY(I)
      GO TO 2
  999 WRITE(6,998)
      IPT=1
      RETURN
      END
```

PUBLICATIONS

The work described in this thesis has formed a basis for the following publications.

1. The determination of sequence of peptides by NMR Spectroscopy.
   J.H. Bradbury, M.W. Crompton, B. Warren.
   Analytical Biochemistry 62 (1974), 310.

2. Proton magnetic resonance spectroscopic studies using paramagnetics of primary and tertiery structure of proteins.
   J.H. Bradbury, L.R. Brown, M.W. Crompton, B. Warren.
   Pure and Applied Chemistry 40, (1974), 83.

3. Determination of the sequence of peptides using paramagnetic probes with NMR Spectroscopy.
   B. Warren and J.H. Bradbury.
   J. Polymer Science. Symposium No. 49. (1975) 65.

4. The determination of the sequence of peptides using relaxation probes.
   J.H. Bradbury and B. Warren.
   Analytical Biochemistry. 78 (1977) 276.

ANALYTICAL BIOCHEMISTRY **78,** 276–278 (1977)

SHORT COMMUNICATIONS

Sequence Determination of Peptides by Measurements of Proton Spin Lattice Relaxation Times

On binding of gadolinium ions to the terminal carboxyl group of a peptide, the spin-lattice relaxation times of the α-C protons are reduced sequentially from the C terminus, thus allowing the sequence determination of submilligram amounts of peptides.

The sequence of peptides has been determined using nuclear magnetic resonance (nmr) spectroscopy by the measurement of (a) shifts of resonances of protons located near the N- and C-terminal residues and caused by changes of pH (1–3); (b) shifts induced by lanthanide ions, Eu^{3+} or Pr^{3+} (2,3); and (c) broadening of resonances resulting from the addition of paramagnetic relaxation probes such as Gd^{3+} and Cu^{2+} (4–6). The first method is limited in application because, even at 300 MHz (2,3), the pH-induced shifts fall off rapidly along the polypeptide chain, such that only 2 or 3 residues can be identified from each end. Methods (b) and (c) require multiple additions of paramagnetic ions to ensure correct assignment in the case of shifting probes and to follow the sequential broadening of resonances in the case of Gd^{3+}. We have found shifting probes to be less useful than broadening probes (4). A further problem with broadening probes is the loss of resolution which becomes apparent as resonances broaden with increasing concentrations of Gd^{3+}. The line broadening with Gd^{3+} results from a predominantly dipolar relaxation mechanism that reduces the spin-spin relaxation time T_2 and also the spin-lattice relaxation time T_1. In this communication, it is shown that measurement of the latter gives an improved method of sequence determination.

Figure 1 shows a typical series of ^1H nmr spectra (offset for clarity) obtained with a Bruker 270-MHz nmr spectrometer using the normal 180-τ-90° pulse sequence for T_1 measurement. The four α-CH_2 proton resonances of tetraglycine are labeled 1–4 beginning from the N terminus in Fig. 1 and are clearly separated. After addition of Gd^{3+} to the solution, the resonances recover from the 180° inversion in the order 4, 3, 2, 1, i.e., the T_1 values have been decreased in that order. A quantitative estimate of the effect is obtained by calculation of T_1 using a computer least-squares analysis from plots of $\ln(S_\infty - S_\tau)$ vs τ, where S_τ and S_∞ are the peak heights for a delay of τ sec and an infinite delay (90° pulse only), respectively. The results obtained for a series of peptides,

ISSN 0003-2697

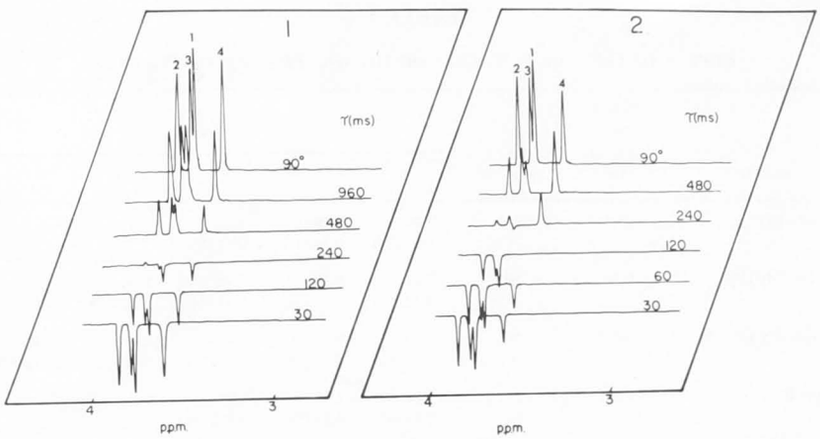


FIG. 1. ^{1}H nmr spectra of the α-CH_2 resonances of tetraglycine (numbered 1 to 4 starting from the N terminus) at 270 MHz using a 180°-τ-90° pulse sequence with various time delays, τ (milliseconds), as shown for (1) tetraglycine (0.10 M) in 6 M guanidine deuterochloride in D_2O at pH meter reading 5.0 and (2) the same system after addition of Gd^{3+} (9×10^{-6} M). The spectra labeled 90° were obtained from the usual 90° pulsing system.

all measured in 6 M guanidine deuterochloride in D_2O (100.0%) at pH 5.0, are given in Table 1. It is seen that the percentage of reduction in T_1 is greatest near the carboxyl groups which binds Gd^{3+} and falls off sequentially along the chain. This allows an unequivocal determination of the sequence in every case except for the last entry in the Table in which there is a considerable amount of binding to the side chain carboxyl group that prevents the correct sequence determination of this peptide [also compare Ref. (6)]. Work is in progress to overcome this difficulty. Another problem arises in this case with the overlap of the α-CH resonances of phenylalanine and alanine.

This method possesses two distinct advantages over our previous studies in which line-broadening methods were used (4–6). First, only a small amount of Gd^{3+} is required, insufficient to cause an appreciable amount of line broadening (see Fig. 1). Second, only one addition of Gd^{3+} is necessary rather than the many additions required previously. By the use of microtubes of 0.05-ml capacity, it is possible to work with submilligram amounts of peptide which can be recovered afterwards as follows. Most of the gadolinium is removed at pH 8–9 as the insoluble hydroxide by centrifugation, and the remainder is removed by complexing with 8-hydroxyquinoline followed by extraction with ethyl acetate. The aqueous phase contains the recovered peptide. The method shows promise as a simple, rapid method for the nondestructive sequencing of small peptides.

TABLE 1

EFFECT OF Gd^{3+} ON T_1 VALUES OF THE α-C PROTONS OF PEPTIDES

Peptide	Peptide concentration (M)	$[Gd^{3+}]^b$ ($M \times 10^6$)	T_1 (sec)[a]				
			1	2	3	4	5
Gly-Gly-Gly	0.15	Nil	1.09	0.78	1.12		
		9.9	0.87 (20)	0.54 (31)	0.51 (55)		
Gly-Gly-Gly-Gly	0.10	Nil	0.61	0.50	0.48	0.64	
		9.0	0.52 (15)	0.37 (26)	0.32 (34)	0.21 (67)	
Gly-Gly-Gly-Gly-Gly	0.025	Nil	0.52	0.46	0.43	0.47	0.42
		1.8	0.48 (9)	0.39 (14)	0.34 (20)	0.32 (32)	0.125 (70)
Ile-Ile-Ile	0.10	Nil	1.15	1.39	1.21		
		18	0.63 (45)	0.63 (55)	0.108 (91)		
Gly-Leu-Gly-Leu	0.05	Nil	0.55	1.02	0.50	1.05	
		5.0	0.34 (38)	0.50 (51)	0.20 (60)	0.20 (81)	
Phe-Asp-Ala-Ser-Val[c]	0.05	Nil	0.72[d]	0.80	0.72[d]	0.83	0.55
		5.5	0.35 (51)	0.121 (85)	0.35 (51)	0.105 (87)	0.043 (92)

[a] Values apply to α-C protons numbered from the N terminus. The numbers given in brackets refer to the percentage reduction in T_1 after addition of Gd^{3+}.

[b] Prepared from $Gd(NO_3)_3\ 5H_2O$ (Koch Light).

[c] T_1 values before and after Gd^{3+} addition for other protons are as follows: β-CH of valine, 0.56 and 0.167 (70); β-CH_2 of aspartic acid, 0.21 and 0.058 (97).

[d] These resonances overlap and the T_1 value is therefore the average for the two nuclei.

ACKNOWLEDGMENT

We wish to thank Dr. Alan Jones, National nmr Centre, Canberra for the 270 MHz nmr spectra and the Australian Research Grants Committee for financial support of this project.

REFERENCES

1. Scheinblatt, M. (1966) *J. Amer. Chem. Soc.* **88,** 2845.
2. Anteunis, M., and Gelan, J. (1973) *J. Amer. Chem. Soc.* **95,** 6502.
3. Anteunis, M., and Callens, R., (1974) *J. Magn. Reson.* **15,** 317.
4. Bradbury, J. H., Crompton, M. W., and Warren, B. (1974) *Anal. Biochem.* **62,** 310.
5. Bradbury, J. H., Brown, L. R., Crompton, M. W., and Warren, B. (1974) *Pure Appl. Chem.* **40,** 83.
6. Warren, B., and Bradbury, J. H., (1975) *J. Poly. Sci.* Part C, **49,** 65.

J. HOWARD BRADBURY
BRIAN WARREN

Chemistry Department
Australian National University
Canberra, A.C.T., Australia
Received May 14, 1976; accepted October 21, 1976

ANALYTICAL BIOCHEMISTRY 62, 310–316 (1974)

Determination of the Sequence of Peptides by PMR Spectroscopy

The outstanding success of the Edman sequenator (1) for the determination of quite long sequences from the N-terminus of peptide chains (see, e.g., Ref. 2), stands in sharp contrast to the dearth of methods available for sequencing from the C-terminus (3). However there are difficulties in the use of the Edman sequenator for small peptides (2) and also in the use of mass spectrometry for sequence determination (4). There is a real need for a method of sequence determination from the C-terminus and for the N-terminal sequencing of small peptides.

Proton magnetic resonance (pmr) spectroscopy can be used to determine the sequences of dipeptides and tripeptides (5) by utilising the changes in the chemical shifts of the α-CH resonances at the N-terminus and C-terminus which result from changes in pH. Very recent work has shown that with the additional resolution available at 300 MHz, tetrapeptides and pentapeptides can also be studied in this way and further gains are possible by the use of lanthanide shift reagents (Pr^{3+}, Eu^{3+}) (6). We have quite independently explored the use of shift reagents (see below) and found them to be much less useful than paramagnetic ions which cause line broadening [Gd^{3+}, Cu^{2+} (7)]. In this paper we report a new method for the sequencing of peptides from the C-terminus and the N-terminus using paramagnetic ions and pmr spectroscopy. The method is nondestructive, has been applied to tri, tetra, and hexapeptides and should be applicable up to a length of ca. 12 residues.

It is a property of polypeptide chains in the random coil state (8) that the average distances from the C-terminal carboxyl group to α-CH protons of successive residues (numbered consecutively from the C-terminus) increases as the number of the residue increases. Similarly, the average distances from the N-terminal amino group to the α-CH protons of successive residues (numbered consecutively from the N-terminus) increases as the number increases. Obviously, the same rule does not apply to the side chain residues which project by different amounts from the backbone of the polypeptide chain. If a paramagnetic ion with a long electronic relaxation time is bound at one or other end of the polypeptide chain, then the dipole–dipole broadening of the α-CH proton resonances along the chain is dependent, amongst other things, on the degree of bind-

ing of the ion at the site and the inverse sixth power of the distances between the bound ion and the α-CH protons (9–12).

If a small amount of the paramagnetic ion is added to a solution of the random coil peptide and the rate of binding is fast at the N- or C-terminus (depending on the ion chosen), then broadening occurs of the α-CH resonance of the nearest amino acid residue. Addition of further amounts of the paramagnetic ion causes increased binding at the terminal site on the polypeptide and sequential broadening of the α-CH resonances (degree of broadening of resonance 1 > resonance 2 > resonance 3, see Fig. 1). The sequential broadening of the α-CH resonances is continued until all have been broadened or until the binding site is saturated with paramagnetic ion. Saturation of the site has not occurred in any system which we have investigated so far. The identity of the various α-CH resonances is obtained from their chemical shifts, spin–spin splitting patterns, and by double resonance experiments.

All experiments were carried out at 100 MHz using a JEOL MH-100 pmr spectrometer. Chemical shifts in Figs. 1 and 2 are recorded for the initial solution in the absence of paramagnetic ions against external tetramethylsilane. Small shifts of all resonances occurred on addition of paramagnetic ions, but in the figures the resonances have been aligned for convenience. Recorded values of pH, are pH meter readings meassured in D_2O. Deuterium oxide was obtained from the Australian Atomic Energy Commission (99.7% D_2O) and Aldrich Chemicals (100.0% D_2O). Gadolinium chloride was prepared from gadolinium oxide (Koch light) and HCl and potassium chromicyanide was prepared by a standard procedure (13). Solutions of Gd^{3+}, $[Cr(CN)_6]^{3-}$ and Cu^{2+} (from cupric chloride) were prepared in D_2O.

In Fig. 1 is shown the result of sequential broadening of the three CH_2 resonances of triglycine from the C-terminus using gadolinium and from the N-terminus with cupric ion. This is a simple case because the CH_2 resonances of glycine have double the intensity of α-CH resonances of other amino acids and also because the latter are split into multiplets (doublets through to quartets). Other peptides which have been successfully sequenced from both the N-terminus using Cu^{2+} and/or $[Cr(CN)_6]^{3-}$ and the C-terminus using Gd^{3+} include Ala Leu Gly and Gly Leu Gly Leu. The sequential broadening of the α-carbon proton resonances of the latter with Gd^{3+} is shown in Fig. 2, with and without decoupling of the β-CH_2 resonances of leucine. In the absence of spin decoupling when Gd^{3+} is added, the α-CH resonances of leucine can be "chemically spin decoupled" (14) from the β-CH_2 groups, causing an increase in the height of α-CH resonances rather than the expected decrease. This could possibly lead to an erroneous determination of the sequence, but the difficulty

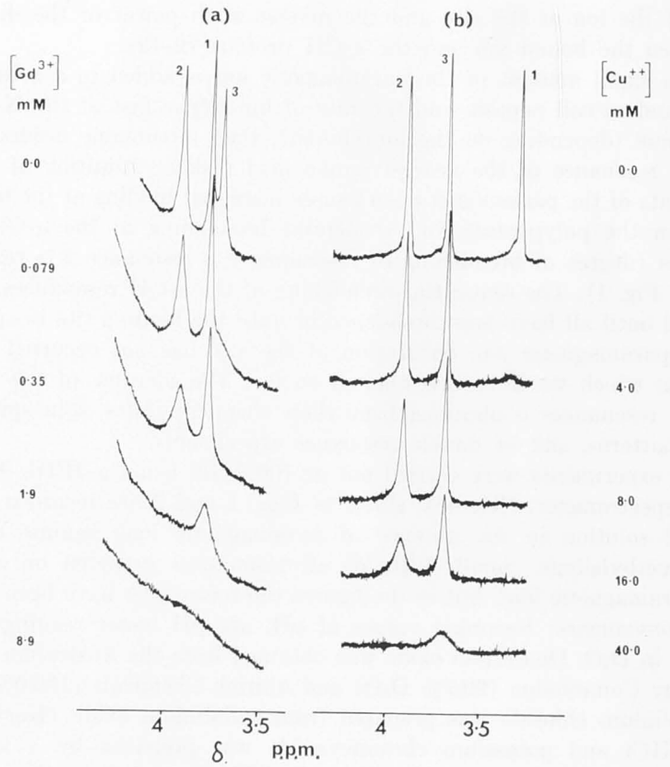


FIG. 1. Pmr spectra at 100 MHz of the methylene protons of triglycine numbered sequentially from the N-terminal end [assignment made by Sheinblatt (5)]. (a) 2% solution of triglycine in 8 M urea at pH 4.3 on addition of Gd^{3+} at the molar concentrations shown, (b) 4% solution of triglycine in D_2O at pH 11, using Cu^{2+}. The differences in the chemical shifts of resonances 1, 2, and 3 in (a) and (b) result from change of pH.

is overcome by spin decoupling as shown in Fig. 2a. This also converts the α-CH leucine triplets into singlets which are readily observed. On addition of Cu^{2+} at pH 11 to hexaglycine, broadening occurs of the N-terminal resonance followed by broadening of the unresolved (at 100 MHz) central CH_2 resonances and finally broadening of the C-terminal CH_2 resonance. Addition of Gd^{3+} at pH 4 to hexaglycine broadens first the C-terminal resonance and then the remaining envelope, which includes the N-terminal resonance. This shows that the paramagnetic ions can effectively broaden the resonances from at least six residues without saturation of the binding site.

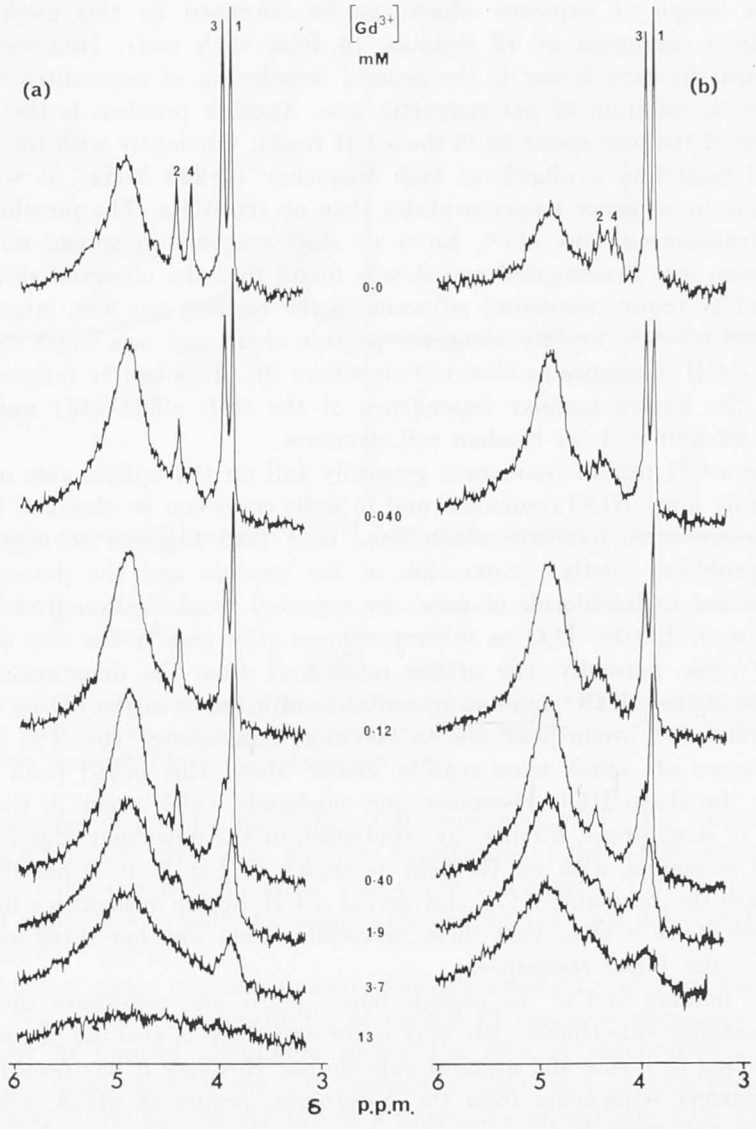


FIG. 2. Pmr spectra at 100 MHz of a 2% solution of Gly Leu Gly Leu in 8 M urea at pH 4.3 in D_2O with the α-carbon proton resonances numbered sequentially from the N-terminal end. The spectra in (b) show the two singlet resonances from glycines 1 and 3 and the overlapping triplet resonances from leucines 2 and 4. The spectra in (a) are obtained by irradiation of both the β-CH_2 resonances of leucines 2 and 4 at 1.64 ppm. The sequential broadening of resonances 4, 3, 2, and 1 is shown on addition of Gd^{3+}.

The length of sequence which can be examined by this method is clearly a minimum of 12 residues (6 from each end). However one practical limiting factor is the general broadening of resonances which occurs on addition of paramagnetic ions. Another problem is the complexity of the pmr spectrum in the α-CH region. Obviously with the additional resolution available at high frequency ($\geq$220 MHz), it will be possible to sequence longer peptides than at 100 MHz. The possible use of paramagnetic ions (Pr^{3+}, Eu^{3+}) as shift reagents to spread out the spectrum was investigated and it was found that the observed shift for the α-CH proton resonance adjacent to the binding site was large, but dropped off very rapidly along the peptide chain and was slight for the third α-CH resonance as observed elsewhere (6). This can be rationalised from the known angular dependence of the shift effect (15) and the great flexibility of the random coil structure.

The α-CH proton resonances generally fall on the upfield side of the normally large HDO resonance, and in some cases can be obscured by it. There are three methods which have been used together to overcome this problem. Firstly, deuteration of the peptide and the denaturant (guanidine hydrochloride or urea) by repeated lyophilisation from D_2O, and use of 100.0% D_2O as solvent reduces very greatly the size of the HDO peak. Secondly, the proton resonances from the denaturant fall downfield from HDO and can by suitable adjustment of the pD be made to form a low broad peak due to exchange broadening (16). The α-CH resonances are much more readily visible above this broad peak than under the sharp HDO resonance and sidebands which occur in the absence of denaturant. Finally, by irradiation of the downfield edge of this broad resonance with an H_2 field as shown in Fig. 3, it is possible to saturate the resonance (17) and reveal α-CH proton resonances hidden beneath it. It is clear that these procedures used together satisfactorily remove the HDO resonance.

The binding of Cu^{2+} to peptide bonds [(18) and references therein] and histidine side chains (19) may occur and there is also the possibility of binding to lysine and arginine side chains. However it has been found that correct sequencing from the N-terminus occurs at pH 8, with no interference from lysine and arginine side chains. Above pH 8, side chain binding of Cu^{2+} on the dipeptides Gly Lys and Ile Arg is appreciable. Histidine binds Cu^{2+} at all pH values and this problem is currently under investigation. The binding of Gd^{3+} on side chain carboxyl groups can be prevented by using conditions where the more weakly acidic side chain carboxyl groups are largely protonated, whilst the α-carboxyl groups are still charged. The apparent pK values of carboxylic acids are increased in 8 M urea (20) by about one unit. We have found that the cor-

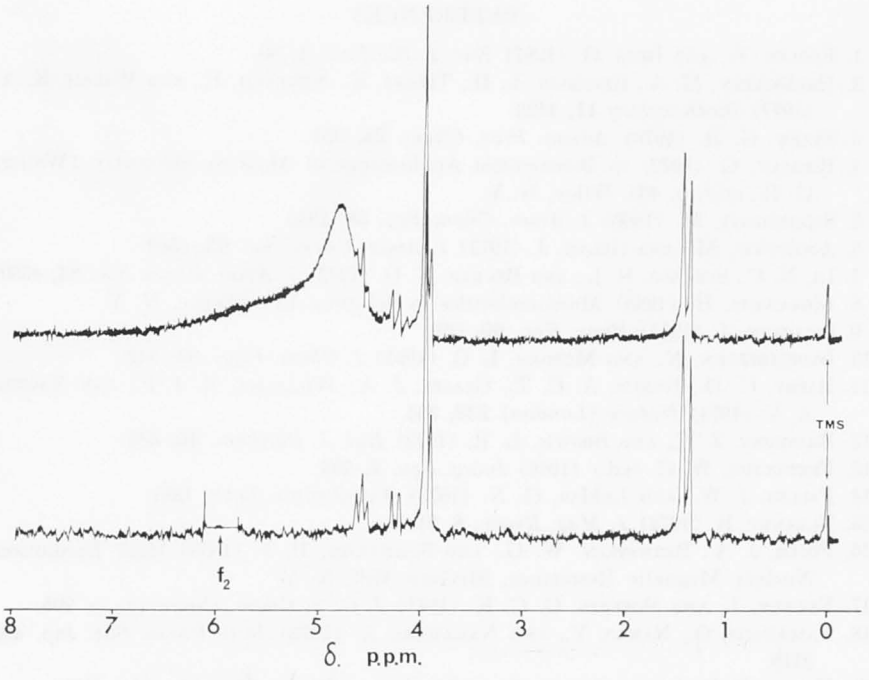


FIG. 3. Pmr spectra at 100 MHz of a 2% solution of Gly Ser Ala in 8 M urea at pH 4.3 in D_2O, with and without irradiation at the frequency f_2. The peptide and urea were deuterated by lyophilisation from D_2O.

rect sequence of broadening of the α-CH resonances of Glu Val Phe occurs at pH 4.3 in 8 M urea, with no more broadening of the γ-CH_2 resonance of Glu, than occurs with other side chain resonances. Thus conditions of pH which are suitable for sequencing are for Gd^{3+} pH 4.3 in 8 M urea and for Cu^{2+} pH 8.

A further major problem concerns the low sensitivity of pmr spectroscopy; the much lower sensitivity of carbon-13 magnetic resonance spectroscopy at natural abundance makes the idea of observation of the selective broadening of the α-C resonances (21) of a peptide even less attractive. However, by the use of a microtube of capacity 0.05 ml and a small amount of spectral accumulation (16 scans) we have been able to obtain an acceptable spectrum of a tripeptide using 0.25 mg of sample. With pulse machines a further tenfold reduction in the amount of material required should be easily possible.

ACKNOWLEDGMENT

We wish to thank the Australian Research Grants Committee for financial support of this project.

REFERENCES

1. Edman, P., and Begg, G. (1967) *Eur. J. Biochem.* **1,** 80.
2. Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972) *Biochemistry* **11,** 4493.
3. Stark, G. R. (1970) *Advan. Prot. Chem.* **24,** 261.
4. Bieman, G. (1972) *in* Biochemical Applications of Mass Spectrometry (Waller, G. R., ed.), p. 405, Wiley, N. Y.
5. Sheinblatt, M. (1966) *J. Amer. Chem. Soc.* **88,** 2845.
6. Anteunis, M., and Gelan, J. (1973) *J. Amer. Chem. Soc.* **95,** 6502.
7. Li, N. C., Scruggs, R. L., and Becker, E. D. (1962) *J. Amer. Chem. Soc.* **84,** 4650.
8. Morawetz, H. (1965) Macromolecules in Solution, Interscience, N. Y.
9. Solomon, I. (1955) *Phys. Rev.* **99,** 559.
10. Bloembergen, N., and Morgan, L. O. (1961) *J. Chem. Phys.* **34,** 842.
11. Barry, C. D., North, A. C. T., Glasel, J. A., Williams, R. J. P., and Xavier, A. V. (1971) *Nature* (*London*) **232,** 236.
12. Bradbury, J. H., and Brown, L. R. (1973) *Eur. J. Biochem.* **40,** 565.
13. Fernelius, W. C. (ed.) (1946) *Inorg. Syn.* **2,** 203.
14. Faller, J. W., and LaMar, G. N. (1973) *Tetrahedron Lett.,* 1381.
15. Bleaney, B. (1972) *J. Mag. Reson.* **8,** 91.
16. Pople, J. A., Bernstein, W. G., and Schneider, H. J. (1959) High Resolution Nuclear Magnetic Resonance, McGraw Hill, N. Y.
17. Feeney, J., and Roberts, G. C. K. (1971) *J. C. S. Chem. Commun.,* p. 205.
18. Yamauchi, O., Nakao, Y., and Nakahara, A. (1973) *Bull. Chem. Soc. Jap.* **46,** 2119.
19. Kruck, T. P. A., and Sarkar, B. (1973) *Can. J. Chem.* **51,** 3549, 3555, 3563.
20. Nelson, C. A., and Hummel, J. P. (1962) *J. Biol. Chem.* **237,** 1567.
21. Bradbury, J. H., and Norton, R. S. (1973) *Biochim. Biophys. Acta* **328,** 10.

J. H. Bradbury
M. W. Crompton
Brian Warren

Department of Chemistry
Australian National University
Canberra, A.C.T., 2600 Australia

Received February 4, 1974; accepted April 25, 1974

Reprinted from **Pure and Applied Chemistry** Vol. 40, Nos. 1–2 (1974).

Pure and Applied Chemistry

The Official Journal of the International Union of Pure and Applied Chemistry

Butterworths

ENGLAND: BUTTERWORTH & CO. (PUBLISHERS) LTD.
LONDON: 88 Kingsway, WC2B 6AB

AUSTRALIA: BUTTERWORTHS PTY. LTD.
SYDNEY: 586 Pacific Highway, Chatswood, NSW 2067
MELBOURNE: 343 Little Collins Street, 3000
BRISBANE: 240 Queen Street, 4000

CANADA: BUTTERWORTH & CO. (CANADA) LTD.
TORONTO: 2265 Midland Avenue, Scarborough, M1P 4S1

NEW ZEALAND: BUTTERWORTHS OF NEW ZEALAND LTD.
WELLINGTON: 26–28 Waring Taylor Street, 1

SOUTH AFRICA: BUTTERWORTH & CO. (SOUTH AFRICA) (PTY) LTD.
DURBAN: 152–154 Gale Street

Printed in Great Britain by Page Bros (Norwich) Ltd., Norwich

PROTON MAGNETIC RESONANCE SPECTROSCOPIC STUDIES USING PARAMAGNETICS OF PRIMARY AND TERTIARY STRUCTURE OF PROTEINS

J. H. Bradbury, Larry R. Brown, M. W. Crompton and Brian Warren

Chemistry Department, Australian National University, Canberra, ACT, Australia

ABSTRACT

On binding a paramagnetic broadening probe at the end of a random coil polypeptide chain, one obtains sequential broadening of the α-CH resonances along the chain and, hence, can determine the sequence of the peptide. The method has been applied to tripeptides (0.25 mg of material being used), tetrapeptides and a hexapeptide under conditions such that gadolinium ions bind at the C-terminus or cupric ions at the N-terminus.

Dimethylation of the lysine residues of lysozyme and observation of the chemical shifts of the six dimethyl proton resonances as a function of pH allows the determination of their p*K*. From studies of selective broadening of the methyl resonances on addition of gadolinium ions and the known structure of the lysozyme–ion complex, it is possible to obtain the p*K* values of the lysine residues of lysozyme.

Approaches to the determination of the structure in solution of proteins using paramagnetic broadening and shifting probes is discussed. Problems concerning the site of ion binding and the assignment of p.m.r. resonances arise when the crystal structure is unknown.

The amount by which a proton magnetic resonance (p.m.r.) line is broadened by a paramagnetic probe with a long electronic relaxation time —e.g. gadolinium(III)—is dependent on the inverse sixth power of the distance between the nucleus giving rise to the resonance and the paramagnetic probe. Furthermore, if the paramagnetic probe is an ion which binds rapidly and reversibly at a specific site, then the width of the resonance line will progressively increase as the concentration of paramagnetic ion is increased until such time as the site is saturated, after which no further broadening will occur.

On the binding of gadolinium ions at low concentration to lysozyme, there is preferential broadening of p.m.r. resonances that arise from protons which are closest to the binding site[1, 2]. The binding site has been shown by x-ray studies[3] to be located between glutamic acid 35 and aspartic acid 52. By the use of difference spectroscopy[4–6], in this case the subtraction of a paramagnetically broadened spectrum from the spectrum of native lysozyme, it has been possible to observe the resonances from the side chains of valine

109 and alanine 110[1,2]. In addition, distance information has been obtained from these studies with broadening probes and also from the use of lanthanide ions with short electronic relaxation times (europium(III), praseodymium(III)), which cause shifts of resonances[2,7]. In this paper three applications are given of the use of paramagnetic ions in peptide and protein chemistry.

1. DETERMINATION OF PRIMARY STRUCTURE (SEQUENCE) OF PEPTIDES

If a peptide is present as a random coil in aqueous solution (which can be achieved by use of a denaturant such as urea or guanidine hydrochloride), then the average distance from one end of the molecule to successive α-CH groups along the peptide chain increases as one proceeds along the chain. Thus, addition of a small amount of a suitable paramagnetic ion which is bound at one end causes most broadening of the α-CH resonance of the nearest amino acid residue, less of the second residue and very little of the third residue[8]. This effect is shown in *Figure 1*, for addition of gadolinium

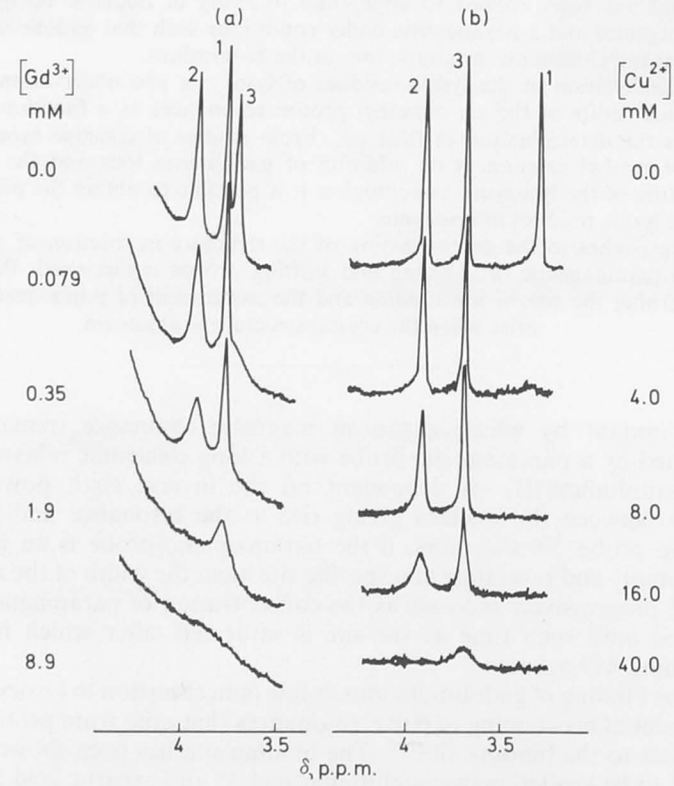


Figure 1. The p.m.r. spectra at 100 MHz of the methylene protons of triglycine numbered sequentially from the N-terminal end (assignment due to Sheinblatt[9]): (a) 2% solution of triglycine in 8 M urea at pH 4.3 on addition of Gd^{3+} at the molar concentrations shown; (b) 4% solution of triglycine in D_2O at pH 11, using Cu^{2+}

ion which binds at the C-terminal carboxylate site and broadens the resonances in the expected sequence 3, 2, 1. Also, cupric ion binds at the N-terminus and causes broadening of residues in the expected sequence 1, 2, 3. It is noted that the differences in the positions of the numbered resonances from (a) to (b) are due to shifts of the resonances which result from the change in pH of the solution[9].

Triglycine is a simple tripeptide to examine because the CH_2 resonances have double the intensity of the α-CH resonances of other amino acids and also because the latter are split into multiplets. A more difficult case is the tetrapeptide Gly–Leu–Gly–Leu, the results of which are given in *Figures 2* and *3*. The N-terminal glycyl CH_2 resonance and the C-terminal leucyl α-CH resonance are readily identified by the manner in which they shift with

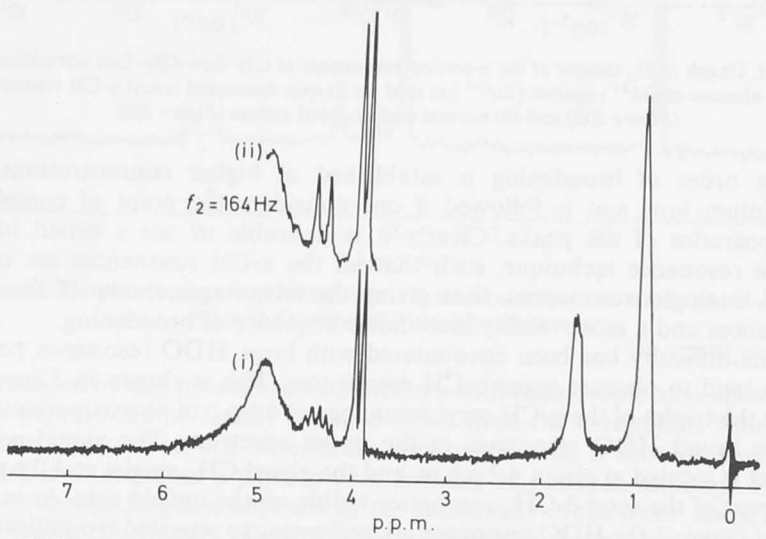


Figure 2. The p.m.r. spectra at 100 MHz of 4% solution of Gly–Leu–Gly–Leu in 8 M urea at pH 4.3: (i) showing two overlapping leucyl α-CH triplets at about 4.3 p.p.m., two CH_2 glycyl resonances at 4 p.p.m. and the β-CH_2 leucyl resonances at 1.6 p.p.m.; (ii) showing the collapse of the α-CH leucyl triplets to singlets owing to irradiation of the β-CH_2 leucyl resonances at 1.6 p.p.m.

change of pH[9]. Irradiation of the β-CH_2 resonances of the leucyl residues at 1.64 p.p.m. causes the collapse of the α-CH triplets into singlets as shown in *Figure 2*. The HDO resonance is small owing to repeated lyophilization of the peptide + urea at pH 4.3 from D_2O and use of 100.0% D_2O (Aldrich Chemical Co.). It is also conveniently broad at this pH because of an intermediate rate of exchange of protons between D_2O and deuterated urea.

The results obtained from the broadening of the resonances on addition of gadolinium ions are shown in *Figure 3*. Under double resonance conditions with all resonances occurring as singlets, the correct order of broadening—4, 3, 2, 1—is obtained. In the absence of decoupling, it is found that the resonance from leucine 2 does not decrease in height (owing to partial decoupling from the β-CH_2) over a range of concentrations for which the resonance from glycine 1 falls off more rapidly. However, it is noted that the

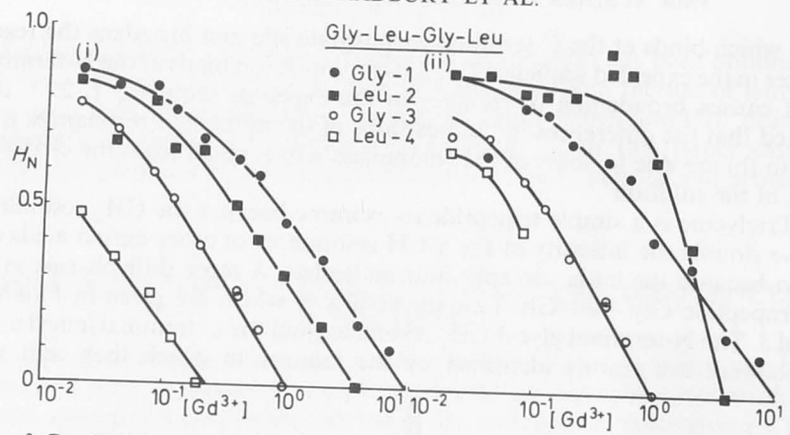


Figure 3. Graph of H_N (height of the α-proton resonances of Gly–Leu–Gly–Leu normalized to 1 in the absence of Gd^{3+}) against $[Gd^{3+}]$ in mM for (i) spin decoupled leucyl α-CH resonances (*Figure 2*(ii)) and (ii) normal undecoupled system (*Figure 2*(i))

correct order of broadening is established at higher concentrations of gadolinium ions and is followed if one considers the point of complete disappearance of the peaks. Clearly it is desirable to use a broad band double resonance technique, such that all the α-CH resonances are converted to singlet resonances, thus giving the advantages shown of sharper resonances and a more readily identifiable sequence of broadening.

Some difficulty has been encountered with large HDO resonance peaks which tend to obscure some α-CH resonances. This is shown in *Figure 4*, where the triplet of the α-CH seryl resonance at 4.6 p.p.m. is partly obscured by the broad HDO resonance in the upper spectrum. The alanyl α-CH quartet is located at about 4.2 p.p.m. and the glycyl CH_2 singlet at 4.0 p.p.m. with part of the seryl β-CH_2 resonance visible on the upfield side. As in the case of *Figure 2*, the HDO resonance is small owing to repeated lyophilization of the peptide + urea from D_2O and broad because of the intermediate rate of exchange of protons between D_2O and deuterated urea. The lower spectrum is obtained by irradiation at the frequency f_2, which causes saturation of the HDO resonance[10].

A second problem concerns the possibility of the binding of gadolinium ions on side chain carboxyl groups and of cupric ions on side chain nitrogen atoms. The former is prevented by working at pH 4.3, since experiments with Glu–Val–Phe have shown that there is no appreciable amount of binding of gadolinium ions to the side chain carboxyl group. This presumably results from the fact that the latter is a weaker acid than the α-carboxyl group; hence, the side chain carboxyl is largely protonated, whereas the α-carboxyl is unprotonated at pH 4.3. Experiments with dipeptides and tripeptides with lysine, arginine and histidine residues in the C-terminal position have shown that at pH 8 in 6 M guanidine hydrochloride there is no appreciable amount of binding to the lysyl or arginyl side chains. Unfortunately there is a considerable degree of binding on the histidyl side chain, and work is in progress on this problem.

A third question concerns the probable length of peptides which can be

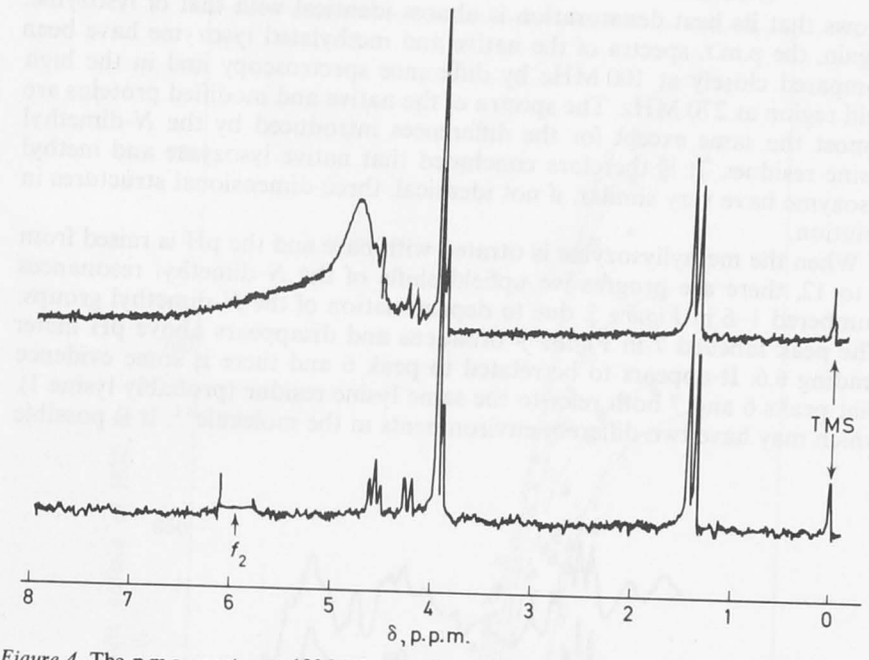


Figure 4. The p.m.r. spectra at 100 MHz of a 2% solution of Gly–Ser–Ala in 8 M urea at pH 4.3 in D_2O, with and without irradiation at f_2

sequenced by this procedure. Experiments with hexaglycine have shown that this peptide can be readily spanned either from the C-terminus with gadolinium ions or from the N-terminus with cupric ions. It should therefore be possible to sequence a peptide with possibly 12 residues, although another limitation on length is imposed by the complexity of the p.m.r. spectrum.

Finally, the amount of peptide required for sequencing has been reduced to about 0.25 mg (for a tripeptide at 100 MHz) by use of a micro-tube of 0.05 ml capacity and a small amount of spectral accumulation[8].

2. DETERMINATION OF DISSOCIATION CONSTANTS OF LYSINE RESIDUES IN LYSOZYME[11]

The ε-amino groups of lysine residues in proteins can be methylated by treatment with formaldehyde and borohydride at 0°[12]. When this procedure is applied to a number of proteins, it is found from amino acid analysis that the extent of dimethylation of lysine residues varies from 66% with bovine serum albumin to 92% for ribonuclease. In lysozyme there are six lysine residues, and the amino acid analysis shows that 4.1 residues are dimethylated and 0.4 residues are N-ε-monomethylated. Since separate proton resonances can be observed from each of the N-dimethyl lysine groups, it is clear that all six lysine residues in lysozyme are reacted.

The *N*-methylated lysozyme retains complete enzymatic activity, and observation of its p.m.r. spectrum as a function of temperature in D_2O

shows that its heat denaturation is almost identical with that of lysozyme. Again, the p.m.r. spectra of the native and methylated lysozyme have been compared closely at 100 MHz by difference spectroscopy and in the high field region at 270 MHz. The spectra of the native and modified proteins are almost the same except for the differences introduced by the *N*-dimethyl lysine residues. It is therefore concluded that native lysozyme and methyl lysozyme have very similar, if not identical, three-dimensional structures in solution.

When the methyllysozyme is titrated with base and the pH is raised from 8 to 12, there are progressive upfield shifts of the *N*-dimethyl resonances numbered 1–6 in *Figure 5*, due to deprotonation of the *N*-dimethyl groups. The peak labelled 7 in *Figure 5* broadens and disappears above pH meter reading 6.6. It appears to be related to peak 6 and there is some evidence that peaks 6 and 7 both refer to the same lysine residue (probably lysine 1), which may have two different environments in the molecule[11]. It is possible

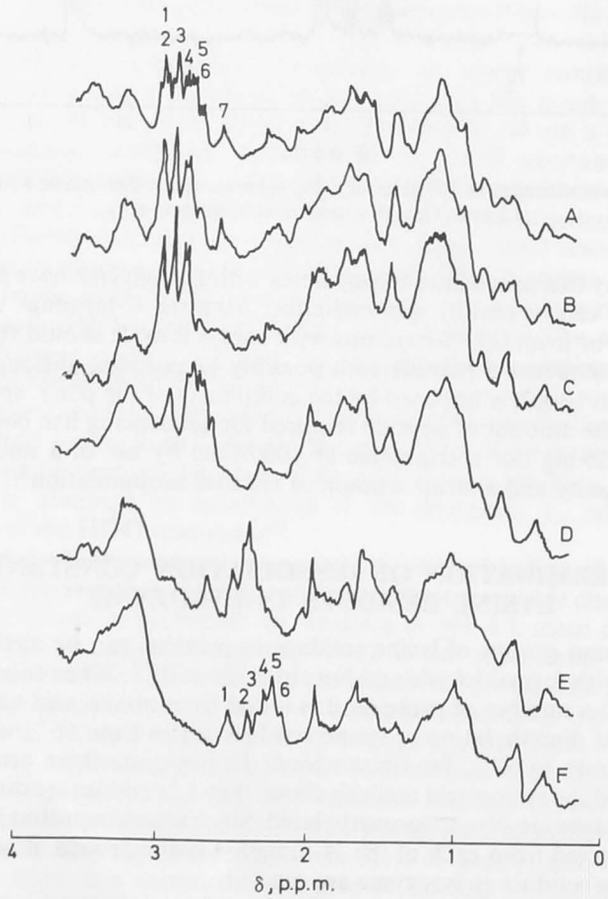


Figure 5. The p.m.r. spectra at 270 MHz of the *N*-methyl resonances of methyllysozyme (2% solution in D_2O) at pH meter reading: A, 3.08; B, 5.08; C, 6.66; D, 9.14; E, 10.17; and F, 10.60

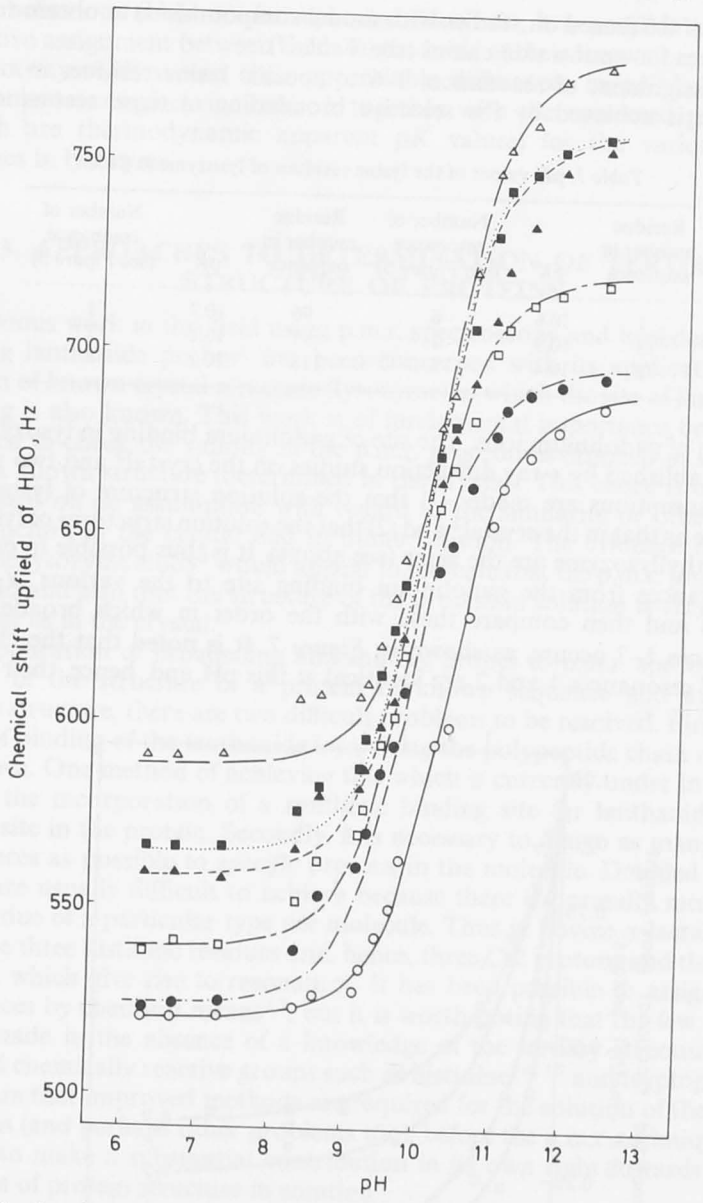


Figure 6. Graphs of chemical shift in hertz upfield from HDO against pH meter reading for *N*-dimethyl resonances of methyllysozyme at 270 MHz and 18°

to follow the peak position of each resonance as a function of pH and, hence, construct the family of titration curves shown in *Figure 6*. These have been analysed as simple, non-interacting ionizations by use of the Henderson–Hasselbach equation and pK values have been determined by a least squares analysis. The range of pK values of 9.6–10.2 refers to *N*-dimethylamino side chains at 18°. A small temperature correction is required to 25° and an

increase of 0.6 (based on studies with model compounds[11]) to obtain reasonable values for amino side chains (see *Table 1*).

The assignment of resonances 1–6 to specific lysine residues in methyllysozyme is achieved by the selective broadening of these resonances on

Table 1. p*K* values of the lysine residues of lysozyme at 25°C

Residue number in sequence	p*K*	Number of resonance (see *Figure 5*)	Residue number in sequence	p*K*	Number of resonance (see *Figure 5*)
1	10.6	6	96	10.7	1
13	10.3	4	97	10.1	2
33	10.4	5	116	10.2	3

addition of gadolinium ions. The site of gadolinium binding in lysozyme has been established by x-ray diffraction studies on the crystal[3] and two reasonable assumptions are made: (1) that the solution structure of lysozyme is the same as that in the crystal[2] and (2) that the solution structures of lysozyme and methyllysozyme are the same (see above). It is thus possible to calculate the distances from the gadolinium binding site to the various *N*-methyl protons and then compare these with the order in which broadening of resonances 1–7 occurs, as shown in *Figure 7*. It is noted that the chemical shifts of resonances 1 and 2 are identical at this pH and, hence, their broad-

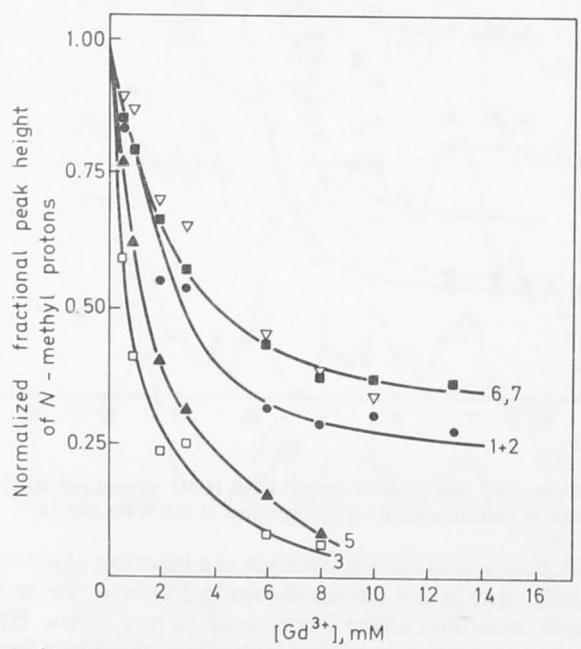


Figure 7. Plot of fractional peak height (normalized against peak 4, which shows very little decrease in height until high gadolinium concentrations and has been used as an internal standard) against $[Gd^{3+}]$ at pH 4.5 in D_2O

ening behaviour could not be separated. However, it is possible to make a tentative assignment between them on the basis of their known environment (from x-ray studies) and their appreciable differences in pK[11]. The final assignments are given in *Table 1*, together with the corrected pK values (which are thermodynamic apparent pK values) for the various lysine residues in lysozyme.

3. APPROACHES TO DETERMINATION OF TERTIARY STRUCTURE OF PROTEINS

Previous work in this field using p.m.r. spectroscopy and broadening and shifting lanthanide probes[2] has been concerned with its application to a protein of known crystal structure (lysozyme) in which the site of lanthanide binding is also known. This work is of fundamental importance because of the need to check the validity of the p.m.r. procedures (applied in solution) using a known structure (determined in the crystal). This check is of course dependent on an assumption with regard to the similarity or otherwise of the structure in the crystal and in dilute solution. The evidence obtained from the lysozyme study[2] would appear to indicate that the p.m.r. procedures are valid and also that the structure of the protein in solution is very closely the same as in the crystal.

In application of broadening and shifting probes to p.m.r. spectroscopic studies of the structure of a protein of known sequence and unknown tertiary structure, there are two difficult problems to be resolved. Firstly, the site(s) of binding of the lanthanide ions along the polypeptide chain needs to be known. One method of achieving this which is currently under investigation is the incorporation of a synthetic binding site for lanthanides at a known site in the protein. Secondly, it is necessary to assign as many p.m.r. resonances as possible to *specific* protons in the molecule. Detailed assignments are usually difficult to achieve because there is normally more than one residue of a particular type per molecule. Thus in bovine α-lactalbumin there are three histidine residues and, hence, three C-2 protons and three C-4 protons which give rise to resonances. It has been possible to assign these resonances by chemical means[13], but it is worth noting that the few assignments made in the absence of a knowledge of the tertiary structure have involved chemically reactive groups such as histidine[14,15] and tryptophan[16]. It appears that improved methods are required for the solution of these two problems (and perhaps other problems too), before the n.m.r. technique will be able to make a substantial contribution in its own right towards determination of protein structure in solution.

ACKNOWLEDGEMENTS

We wish to thank Drs R. E. Richards and I. D. Campbell for providing access to the facility of the 270 MHz p.m.r. spectrometer at Oxford University and Professor D. C. Phillips for provision of laboratory facilities during a period of sabbatical leave by J.H.B. in 1972. The Australian Research Grants Committee is thanked for the provision of financial support for this work.

REFERENCES

[1] I. D. Campbell, C. M. Dobson, R. J. P. Williams and A. V. Xavier, *J. Magn. Resonance*, **11**, 172 (1973).
[2] I. D. Campbell, C. M. Dobson, R. J. P. Williams and A. V. Xavier. *Ann. NY Acad. Sci.* **222**, 163 (1973).
[3] C. C. F. Blake and M. A. Rabstein, personal communication.
[4] B. Bak, E. J. Pedersen and F. Sundby, *J. Biol. Chem.* **242**, 2637 (1967).
[5] N. L. R. King and J. H. Bradbury, *Nature* (*Lond.*) **229**, 404 (1971).
[6] J. H. Bradbury and N. L. R. King, *Austral. J. Chem.* **25**, 209 (1972).
[7] C. D. Barry, A. C. T. North, J. A. Glasel, R. J. P. Williams and A. V. Xavier, *Nature* (*Lond.*) **232**, 236 (1971).
[8] J. H. Bradbury, M. W. Crompton and Brian Warren, *Analyt. Biochem.* **62**, 310 (1974).
[9] M. Sheinblatt, *J. Amer. Chem. Soc.* **88**, 2845 (1966).
[10] J. Feeney and G. C. K. Roberts, *Chem. Commun.* 205 (1971).
[11] J. H. Bradbury and Larry R. Brown, *Europ. J. Biochem.* **40**, 565 (1973).
[12] G. E. Means and R. E. Feeney, *Biochemistry*, **7**, 2192 (1968).
[13] R. S. Norton, *Ph.D. Thesis*, Australian National University (1974).
[14] D. H. Meadows, O. Jardetzky, R. M. Epand, H. Rüterjans and H. A. Scheraga, *Proc. Nat. Acad. Sci., Wash.* **60**, 766 (1968).
[15] J. H. Bradbury and B. E. Chapman, *Biochem. Biophys. Res. Commun.* **49**, 891 (1972).
[16] J. D. Glickson, W. D. Phillips and J. A. Rupley, *J. Amer. Chem. Soc.* **93**, 4031 (1971).

J. POLYMER SCI.: Symposium No. 49, 65-74 (1975)

DETERMINATION OF THE SEQUENCE OF PEPTIDES USING PARAMAGNETIC PROBES WITH NMR SPECTROSCOPY

BRIAN WARREN and J. H. BRADBURY

Chemistry Department, Australian National University, Canberra, A.C.T., Australia

SYNOPSIS

The paramagnetic ion Gd^{3+} binds to the α-carboxyl group of a peptide and broadens the resonances from the ^{13}C and ^{1}H nuclei of the backbone sequentially along the chain. The sequential broadening of both carbonyl and C^{α} nuclei of PheAspAlaSerVal have been observed in 6 M guanidine hydrochloride at pH 4.0. Sequential broadening from the N-terminus is also possible using Cu^{2+} at pH $\geqslant 8$.

Broadening experiments using triglycine and either Gd^{3+} or Cu^{2+} show that it is more difficult to broaden ^{13}C than ^{1}H resonances on the same amino acid residue. The ratio of $[Gd^{3+}]_{CMR}/[Gd^{3+}]_{PMR}$ is 12 ± 5, which agrees reasonably well with the theoretical value of 16. The ratio $[Cu^{2+}]_{CMR}/[Cu^{2+}]_{PMR}$ is 3 ± 2.

The relative advantages of proton magnetic resonance (PMR) and carbon-13 magnetic resonance (CMR) spectroscopy for this application are discussed. It is concluded that sequence determinations of 20 mg samples of peptides, ranging in size up to 7-9 residues, should be possible using CMR spectroscopy. The main advantage of PMR spectroscopy is the much smaller amounts of material required, but other problems preclude its use at this time except for very simple peptides.

INTRODUCTION

The use of paramagnetic probes in conjunction with nuclear magnetic resonance (NMR) spectroscopy has become of considerable importance in chemistry [1, 2] and biology [3]. An important quantitative application which is possible under certain conditions [1, 3, 4] is the determination of the distance (r) between a paramagnetic species and the nucleus of interest. One set of limiting conditions can be imposed by the use of paramagnetic ions with long electronic relaxation times and for which contact (through bonds) interactions are negligible [3, 4]. In this case the Solomon-Bloembergen equations [2-4] indicate that the paramagnetic contribution to the spin-spin relaxation rate $(1/T_{2M}) \alpha^{al} 1/r^6$. Since the total line width of a resonance is made up of various contributions, including that of $(1/T_{2M})$ [5], it is possible to obtain ratios of distances from the line widths of NMR resonances [6].

A qualitative application of the same principle is that the amount of broadening of the resonances from two or more nuclei which are near a bound paramagnetic species will increase with their proximity to it. Thus, on increasing the concentration of paramagnetic ions in solution, and hence also the amount of binding at a site, the resonance from the nearest nucleus will broaden first and that from the furthest nucleus last. This procedure has been used to assign the six $-N(CH_3)_2$ proton resonances of the lysine residues of lysozyme [7, 8] and also to determine the sequence of amino acids in a peptide [8, 9]. Until recently we have used PMR spectroscopy, but because of the much greater range of chemical shifts available with carbon-13, studies have now been made using this nucleus. The present paper describes these studies and gives a comparison between the relative merits of the two nuclei.

EXPERIMENTAL

The various peptides used in this work were obtained from the Fox Chemical Co. and Mann Research Laboratories, and their purity was checked by PMR and CMR spectroscopy before use. For ^1H studies a Jeol 100 MHz PMR spectrometer (MH-100) was used, and for ^{13}C studies a Brüker HX-90 spectrometer operating at 22.63 MHz with proton-noise decoupling. The pulse width and the delay time were varied in order to obtain optimal signal averaging over a constant time interval. The amount of spectral accumulation varied from about 200-4000 scans. One-ml samples of 10% (w/v) solutions of peptides were examined in 10-mm diameter sample tubes in 6 M guanidine hydrochloride (GuCl) in water at the pH indicated. A small amount of D_2O was added to the solution to serve as a lock signal for the spectrometer. Chemical shifts are recorded in ppm from the ^{13}C resonance of GuCl (upfield shifts shown positive) which was shown to occur 159.4 ppm downfield from external tetramethylsilane (TMS), over the pH range from 4 to 10. Peak positions were obtained from a computer print-out and are accurate to about ±0.1 ppm.

RESULTS

Figure 1 shows the CMR spectra of a pentapeptide, PheAspAlaSerVal, and various other peptides containing similar residues whose resonances have been used to identify those of the pentapeptide. The exact resonance positions and the identities of the resonances are given in Table I. The assignment of the aromatic carbon resonances is based on previous studies [10]. The C^β resonance of aspartic acid is poorly defined in both the pentapeptide and GlyAspSer, but its position has been confirmed in another tripeptide ProPheAsp, the spectrum of which is not shown. The position of all the C^α resonances is readily obtained by comparison of the line positions of the model peptides in Table I.

The assignment is more difficult for the carbonyl and carboxyl CMR resonances which are usually much smaller than the C^α resonances, because of a

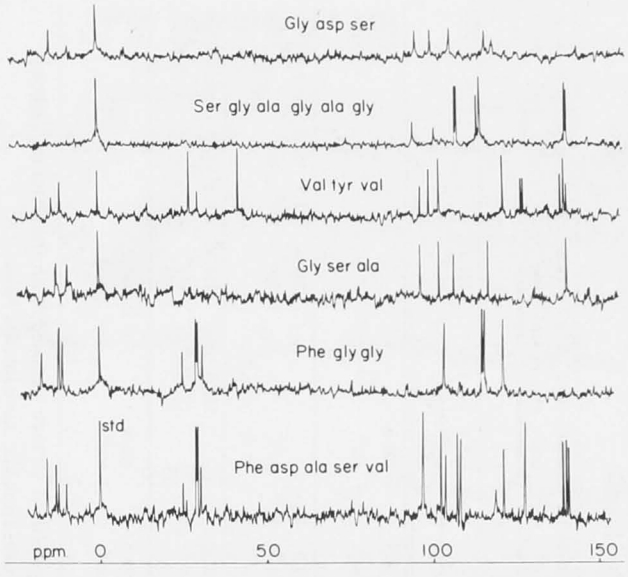


FIG. 1. CMR spectra of peptides in 6 M GuCl at pH-4.0 in water containing 5% D_2O. Chemical shifts are in ppm upfield from the GuCl resonance.

smaller or zero nuclear Overhauser enhancement (NOE) [11]. In addition, their T_1 values are usually considerably larger [12], which results in slower relaxation to equilibrium of their nuclei after pulsing. Thus if the pulse width is large (an intense pulse) and the recycle time is short, the carbonyl resonances may not be observed at all, as is the case for the hexapeptide in Figure 1. Inspection of Table I shows that the positions of the carbonyl resonances of Phe, Asp, and Ser can be determined by comparison of line positions. The valine carboxyl resonance is not observed in the spectrum in Figure 1, but was observed in another spectrum of the pentapeptide (see Fig. 2, resonance A) and its position is confirmed by that of ValTyrVal in Figure 1. The assignment of resonance B to the carbonyl resonance of alanine is then obtained for the pentapeptide by difference.

As would be expected, the peak positions of the resonances are dependent on their positions in the peptide (whether N-terminal, internal, or C-terminal) and undoubtedly (to a lesser extent) on the identity of the adjacent residue ([13] and references therein). The former dependence is shown by the results in Table II for the carbonyl and the C^{α} resonances. In general there is a downfield shift of each of those resonances from the N-terminal to internal to C-terminal positions. However, this trend is most marked for the carbonyl and C^{α} resonances of glycine in Table II, is not observed for the C^{α} resonances of glycine in triglycine (Fig. 3), and is reversed for the internal C^{α} resonance of serine in Table II. An indication of the variability of the peak positions which results from different neighboring residues is also given by the multiple entries in Table II. The average variability between pairs of entries for C^{α} resonances is about 0.3 ppm, whereas that for C=O resonances (excluding carboxyl groups because of the possibility of their variable degree of charging at pH 4.0) is 0.6 ppm.

TABLE I

Carbon-13 Chemical Shifts (ppm upfield from GuCl) of Resonances of Peptides in 6 M GuCl at pH-4.0[a] (see Figs. 1 and 2)

PheAspAlaSerVal	-20.2 Val CO	-16.4 Ala CO	-13.6 Asp CO	-13.1 Ser CO	-10.9 Phe CO	24.5 Phe C^{γ}	28.6 Phe C^{ε}	28.9 Phe C^{δ}	30.1 Phe C^{ζ}	96.8 Val C^{α} Ser C^{β}	102.3 Ser^{α}	103.6 Phe^{α}	106.8 Asp^{α}	108.2 Ala^{α}	119.7 Asp^{β}	121.2 Phe^{β}	127.7 Val^{β}	139.3 $Val^{\gamma 1}$	140.5 $Val^{\gamma 2}$	141.2 Ala^{β}
	A[b]	B	C	D	E					F	G	H	I	J	K	L	M	N	O	P
PheGlyGly		-17.8 Gly COO^-		-12.2 Gly CO	-11.4 Phe CO							103.4 Phe^{α}	114.6 Gly^{α} COO^-	115.3 Gly^{α}		121.3 Phe^{β}				
GlySerAla	Ala COO^- not obs[c]			-12.7 Ser CO	-9.5 Gly CO					96.5 Ser^{β}	102.4 Ser^{α}		106.8 Ala^{α}		117.1 Gly^{α}					140.8 Ala^{β}
ValTyrVal		-18.4 Val COO^-	-14.0 Val CO		-11.6 Tyr CO	Tyr C^{ζ} not obs[c]	27.6 Tyr C^{δ}	30.0 Tyr C^{γ}	42.5 Tyr C^{ε}	96.9 Val C^{α} COO^-	99.5 Val^{α}	102.5 Tyr^{α}				121.8 Tyr^{β}	127.4 Val^{β}	127.8 Val^{β}	139.2/140.2/141.0 $Val^{\gamma 1}{}^{\gamma 2}$		
SerGlyAlaGlyAlaGly	Carbonyl resonances not observed									95.5 Ser^{β}	102.0 Ser^{α}			108.4/108.9 Ala^{α}		114.9[d] Gly^{α} COO^-	116.0[d] Gly^{α}		141.7/142.1 Ala^{β}		
GlyAspSer	Ser COO^- not obs[c]		-14.0 Asp CO		-8.4 Gly CO					95.9 Ser^{β}	100.4 Ser^{α}		106.4 Asp^{α}		117.0 Gly^{α}	119.2 Asp^{β}				

[a] Pulse width 2.5 μsec and recycle time 0.4 sec, except for the hexapeptide where a pulse width of 12 μsec was used. With the latter no carbonyl resonances were observed because of their larger T_1 values and hence inability to recover from the more intense pulse.

[b] These letters refer to the resonances indicated in Figure 2.

[c] These resonances were not observed, probably because of their longer T_1 values (see footnote[a]).

[d] The α-C resonance of the C-terminal glycine residue is distinguished because its intensity is about one-half that of the α-C resonance from the two internal glycine residues.

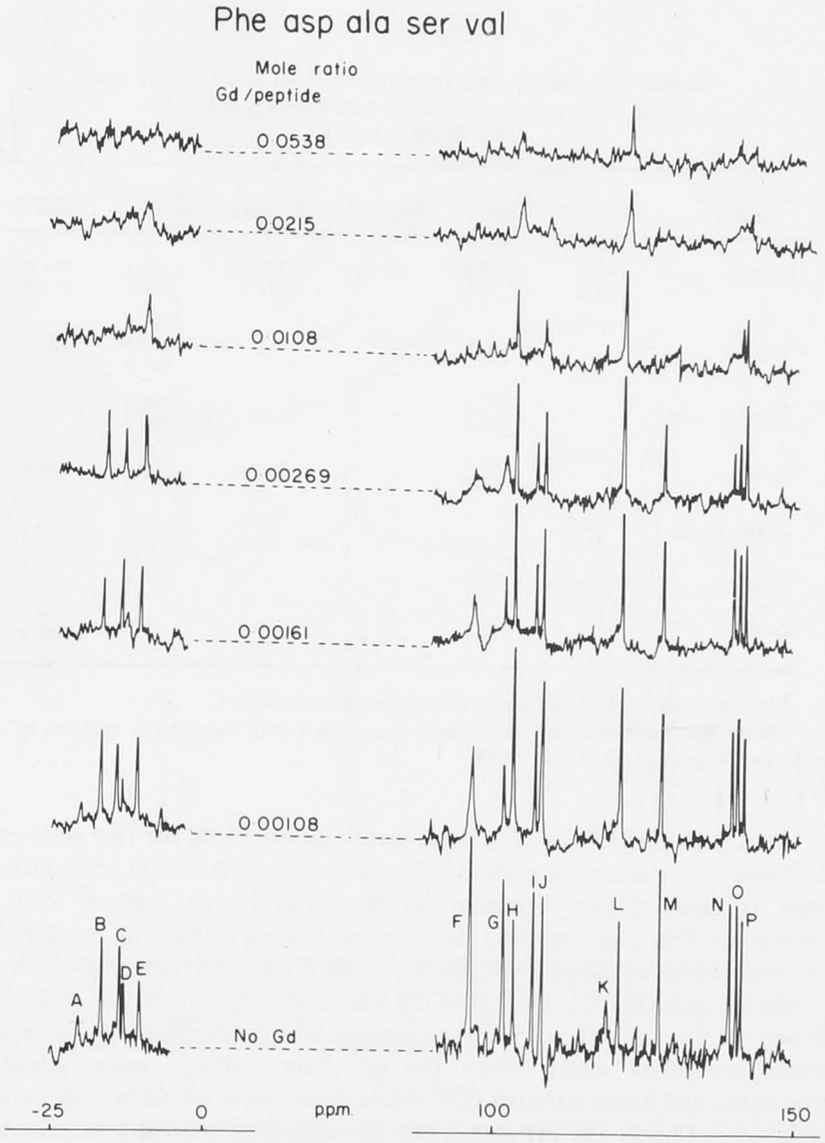


FIG. 2. CMR spectra of a 0.186 M solution of the pentapeptide PheAspAlaSerVal in the presence of increasing amounts of Gd^{3+} in 6 M GuCl at pH-4.0. The assignments of the resonances labeled A to P are based on the data in Fig. 1 and are tabulated in Table I. Chemical shifts are in ppm upfield from GuCl.

The broadening of the carbonyl and C^{α} resonances of PheAspAlaSerVal on addition of Gd^{3+} is shown in Figure 2. If one compares this figure with the assignments in Table I it is seen that the order of broadening of the C=O resonances is that expected from binding of Gd^{3+} at the C-terminus, viz., Val (A) > Ser (D) > Ala (B) > Asp (C) > Phe (E). However the broadening of the C^{α} resonances is more complex. The order of broadening is resonance F (composite

TABLE II

Carbon-13 Chemical Shifts (ppm upfield from GuCl) of Carbonyl and C^{α} Resonances of Amino Acid Residues in Peptides in 6 M GuCl at pH 4.0[a]

Amino acid residue	C=O resonances			C^{α} resonances		
	N-terminal	internal	C-terminal[b]	N-terminal	internal	C-terminal
Glycine	-9.5 -8.4	-12.2	-17.8	117.1 117.0	115.3 116.0	114.6 114.9
Alanine	-	-16.4	-	-	108.2 108.4 108.9	106.8
Aspartic acid	-	-13.6 -14.0	-	-	106.8 106.4	-
Serine	-	-13.1 -12.7	-	102.0	102.3 102.4	100.4
Phenylalanine	-10.9 -11.4	-	-	103.6 103.4	-	-
Tyrosine	-	-11.6	-	-	102.5	-
Valine	-	-	-20.2 -18.4	99.5	-	96.8 96.9

[a]Entries in this table are obtained from inspection of Table I.
[b]These line positions may be somewhat variable due to incomplete charging of carboxyl group at pH 4 in 6 M GuCl.

of Val C^{α} and Ser C^{β} which reduces greatly in height on the first addition of Gd^{3+} and then much more slowly) > Ser (G) > Asp (I) > Ala (J) > Phe (H). The order is inverted here between the third and fourth residues from the C-terminus. This is undoubtedly due to some binding of Gd^{3+} on the side-chain carboxyl groups of aspartic acid, indicated by the rapid broadening of the Asp C^{β} resonance K (Fig. 2). The pH of the broadening experiment was deliberately chosen on the basis of previous experiments with GluValPhe [9], in order to obtain conditions under which the side-chain carboxyl group would be protonated and hence not bind Gd^{3+}, while the α-COOH would be only partially protonated. Clearly the pH used is not low enough to prevent some side-chain binding. Further experiments are in progress on this matter. The fact that the correct order of broadening is obtained with the C=O but not with the C^{α} resonances can be explained by assuming relatively weak binding on the aspartyl carboxyl, which is nevertheless strong enough to broaden its C^{α} resonance (before that of the Ala C^{α} resonance) but not its more distant C=O resonance (before that of the Ala C=O resonance).

Sequential broadening of the pentapeptide in 6 M GuCl at pH-8 using Cu^{2+} from the N-terminus, gives the correct order of broadening of the first three residues, but Cu^{2+} is much less effective than Gd^{3+} for subsequent residues. This is related to the fact that broadening with Cu^{2+} is largely due to contact (through bond) interactions and also to the complexity of the binding of Cu^{2+} to peptides [14]. These matters will be considered in a later publication.

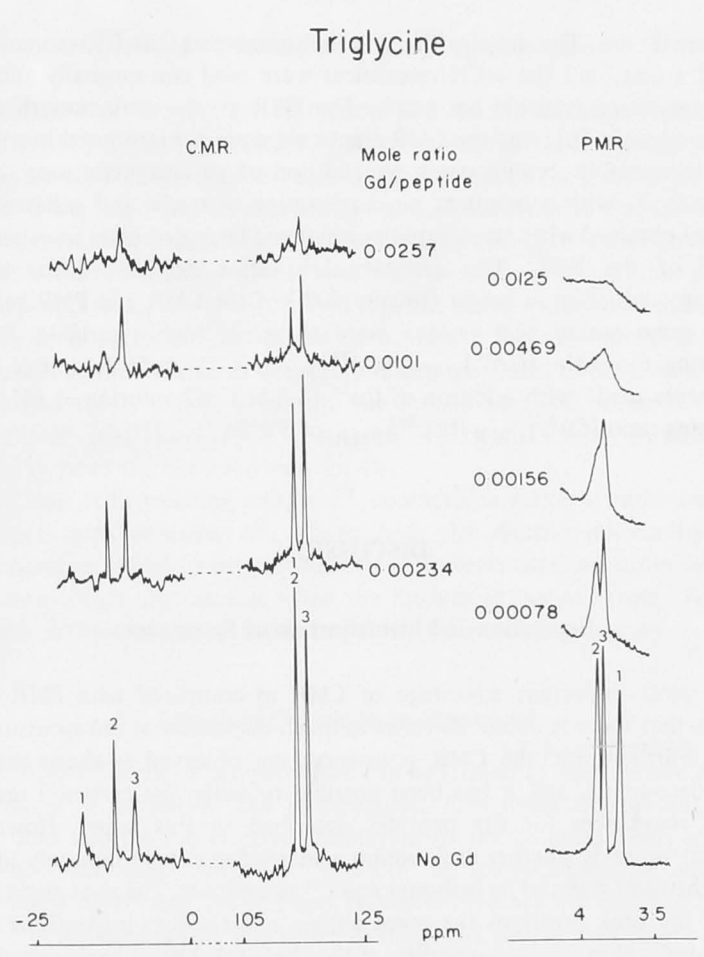


FIG. 3. Comparison of PMR spectra (ppm from TMS) and CMR spectra (ppm upfield from GuCl) of 0.512 M solutions of triglycine in 6 M GuCl at pH-4.0 in D_2O on addition of Gd^{3+}. The resonances within any one group, i.e., C=O (0 to 25 ppm), C^{α} (105-125 ppm), and CH_2 (3.5-4 ppm), are labeled 1, 2, and 3 numbering from the C-terminus.

In Figure 3 the results are given of parallel experiments on the broadening of the PMR and CMR resonances of triglycine on the addition of Gd^{3+} at pH 4. It is noted that while the C=O resonances are well separated, the C-terminal and central C^{α} resonances of the tripeptide have the same chemical shifts. The positions of the resonances in 6 M GuCl have been assigned by following the shifts of the resonances as a function of pH. The broadening of the C=O and C^{α} resonances follows the expected sequence. The assignment of the PMR resonances was originally due to Sheinblatt [15] and the broadening of these resonances follows the expected sequence as described previously [9].

In an attempt to obtain a semiquantitative estimate of the ratio of the concentrations of paramagnetic ions required to broaden a CMR resonance as compared with a PMR resonance on the same amino acid residue, the fractional peak heights of the resonances have been graphed against the concentration of

paramagnetic ion. The data for the carbonyl resonances, the C^{α} resonances from single C atoms, and the α-CH resonances were used and normally about eight data points were available per graph. The PMR graphs were smooth curves as found previously [8]; but the CMR graphs were much more variable with initial height increases in certain cases on addition of paramagnetic ions (due to a decrease in T_1 with consequent rapid relaxation of nuclei and enhancement of the signal obtained when using rapid pulsing) and large decreases in others due to removal of the NOE. The concentration ratios determined for a certain percentage reduction in height (usually 50%) of the CMR and PMR resonances on the same amino acid residue were therefore highly variable. For Gd^{3+} broadening the ratio $[Gd^{3+}]_{CMR}/[Gd^{3+}]_{PMR}$ is 12 ± 5. A similar series of experiments made with addition of Cu^{2+} to 6 M GuCl solution at pH 8 gives a broadening ratio $[Cu^{2+}]_{CMR}/[Cu^{2+}]_{PMR}$ of 3 ± 2.

DISCUSSION

Separation and Identification of Resonances

The most important advantage of CMR as compared with PMR spectroscopy, is that there is about 20 times as much dispersion of the spectrum in the former. Furthermore, the CMR resonances are observed as sharp singlets by proton decoupling, and it has been possible to assign the carbonyl resonances and C^{α} resonances for the peptides described in this paper. However the important point is whether each amino acid residue can be uniquely identified by the chemical shifts of its carbonyl and C^{α} resonances. The data given in Table II show the peak positions for seven amino acid residues in peptides and also gives an indication of the variability of the chemical shift of peaks which results from different neighboring residues. The total chemical shift range for carbonyl and carboxyl resonances is about 12 ppm and the sequence variability is about 0.6 ppm. The corresponding figures for the C^{α} resonances are about 21 ppm and 0.3 ppm. At first sight it would appear that all the amino acid residues may be separable from one another, but more complete compilations of chemical shift data for amino acid residues in peptides in D_2O and dimethylsulfoxide [13] indicate the likelihood of considerable ambiguity between certain residues using our solvent system. On the other hand, it has been possible to assign all the carbon resonances in the decapeptide gramicidin S-A [16].

By comparison, the PMR method suffers from the disadvantages that the α-CH region of the spectrum falls over a rather narrow range, which is overlapped on the downfield end by the HDO resonance, with all resonances, except those from glycine, occurring as doublets through to quartets [9]. However, methods are available for reducing the size of the HDO resonance (see below) and it can be completely eliminated by a double resonance technique [9]. Also it may be possible to collapse all the αCH resonances to singlets by means of broad-band decoupling techniques.

Relative Broadening of ^{1}H and ^{13}C Resonances

Examination of the dipolar relaxation term in the Solomon-Bloembergen equation [3, 17] indicates that $1/T_{2M}$ is proportional to γ^2 (where γ is the magnetogyric ratio of the nucleus) and is independent of the Larmor precession frequency, providing there is rapid motion as occurs with small molecules. Since the paramagnetic line broadening is proportional to $1/T_{2M}$, it would appear that the relative line broadening of ^{1}H and ^{13}C resonances on addition of Gd^{3+} should be proportional to $\gamma_H{}^2/\gamma_C{}^2 = 16$. For the simple case shown in Figure 3, where the mole ratio of Gd^{3+} added to peptide is less than about 0.05, the concentration of Gd^{3+} added to the solution is proportional to the concentration of the bound Gd^{3+}, which is causing the broadening. The observed ratio of concentrations $[Gd^{3+}]_{CMR}/[Gd^{3+}]_{PMR}$ is 12 ± 5, which is within experimental error of the expected value of 16.

In the case of broadening using Cu^{2+}, contact interaction is important [17]. The contact term contains A^2, where A is the electron-nuclear hyperfine coupling constant which is proportional to γ. However there are other factors in the equation which also change when the nucleus is changed from ^{1}H to ^{13}C [18], and it is therefore not possible to predict the result from theory.

Length of Peptide to be Sequenced

The ultimate length of peptide which can be treated by this method depends on the range of action of the paramagnetic ions when the binding site is saturated and also on the effective resolution of the complex NMR spectrum [8, 9]. The CMR results with Gd^{3+} and Cu^{2+} on the pentapeptide indicate an effective range of 4-5 and 3-4 residues, respectively. Since it is easier to broaden ^{1}H than ^{13}C resonances (see preceding section), the former may have a slightly longer range of action. On the other hand, it is much easier to resolve the CMR as compared with the PMR spectrum and currently this greatly limits the use of the latter. The earlier estimate of 12 residues (6 from each end) [8, 9] which was based on PMR studies of hexaglycine may be too large and a value of 7-9 seems to be more realistic.

Quantity of Peptide

CMR spectroscopy is intrinsically much less sensitive than PMR spectroscopy as shown by the fact that about 100 mg of peptide is used in the CMR work, compared with a minimum of 0.2 mg of material in PMR studies, with a continuous wave spectrometer [9]. By working at lower concentrations and with more spectral accumulation it should be possible to reduce the amount of peptide needed to the order of 20 mg and 0.04 mg, respectively. This shows clearly the greatest single advantage of PMR as compared with CMR spectroscopy.

Sample Preparation and Solvents

The solvent of choice for the CMR studies is water containing a denaturant (6 M GuCl) to ensure that the peptide is present in the random coil form. The carbon-13 resonance from GuCl is very useful as an internal standard, since its chemical shift compared with external tetramethylsilane is independent of pH over a wide range and remains unaffected by the presence of paramagnetic ions. With PMR spectroscopy, the solvent consists of D_2O and a denaturant (GuCl or urea). It is important to minimize the size of the HDO resonance by repeated lyophilization from D_2O of the peptide and the denaturant, using 100.0% D_2O and double resonance techniques [9].

The authors wish to thank Mr. M.W. Crompton for useful discussions, the staff of H.B. Selby and Co., Melbourne, for assistance in obtaining the CMR spectra, and the Australian Research Grants Committee for financial suppport.

REFERENCES

[1] J. Reuben, in *Progress in Nuclear Magnetic Resonance Spectroscopy*, Vol. 9, Part 1, J. W. Emsley, J. Feeney, and L. H. Sutcliffe, Eds., Pergamon Press, Oxford, 1973.
[2] G.N. LaMar, W. DeW. Horrocks, and R.H. Holm, *NMR of Paramagnetic Molecules*, Academic Press, New York., 1973.
[3] R.A. Dwek, *Nuclear Magnetic Resonance in Biochemistry*, Clarendon Press, Oxford, 1973.
[4] C.D. Barry, J.A. Glasel, R.J.P. Williams, and A.V. Xavier, *J. Mol. Biol., 84,* 471 (1974).
[5] J.H. Bradbury, B. E. Chapman, and N.L.R. King, *Int. J. Protein Res., 3,* 351 (1971).
[6] C.D. Barry, A.C.T. North, J. A. Glasel, R.J.P. Williams, and A.V. Xavier, *Nature, 231,* 236 (1971).
[7] J.H. Bradbury and L.R. Brown, *Eur. J. Biochem., 40,* 565 (1973).
[8] J.H. Bradbury, L.R. Brown, M.W. Crompton, and B. Warren, *Pure Appl. Chem.*, in press.
[9] J.H. Bradbury and M.W. Crompton, and B. Warren, *Anal. Biochem., 62,* 310 (1974).
[10] J.H. Bradbury and R.S. Norton, *Biochim. Biphys. Acta, 328,* 10 (1973).
[11] J.H. Noggle and R.E. Schirmer, *The Nuclear Overhauser Effect*, Academic Press, New York, 1971.
[12] A. Allerhand and R.A. Komoroski, *J. Amer. Chem. Soc., 95,* 8228 (1973).
[13] C. Grathwohl and K. Wüthrich, *J. Mag. Reson., 13,* 217 (1974).
[14] R.P. Agarwal and D.D. Perrin, *Pure Appl. Chem., 34,* 387 (1972).
[15] M. Sheinblatt, *J. Amer. Chem. Soc., 88,* 2845 (1966).
[16] J.A. Sogn, L.C. Craig, and W.A. Gibbons, *J. Amer. Chem. Soc., 96,* 3306 (1974).
[17] T.J. Swift, in *NMR of Paramagnetic Molecules*, G.N. LaMar, W. DeW. Horrocks, and R.H. Holm, Eds., Academic Press, New York, 1973, ch. 2.
[18] B.R. McGarvey and R.J. Kurland, in *NMR of Paramagnetic Molecules*, G.N. LaMar, W. DeW. Horrocks, and R.H. Holm, Eds., Academic Press, New York, 1973. ch. 14.